

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 12:59:49 ; Search time 9150.23 Seconds
(without alignments)
10942.062 Million cell updates/sec

Title: US-09-303-232-1_COPY_372_2681

Perfect score: 2310

Sequence: 1 atgaaatgcacactgaa.....caccacatatattgtctctcg 2310

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_or.*
- 21: em_ov.*
- 22: em_ph.*
- 23: em_pl.*
- 24: em_ro.*
- 25: em_sts.*
- 26: em_un.*
- 27: em_vi.*
- 28: em_hum.*
- 29: em_hum.*
- 30: em_hum.*
- 31: em_hum.*
- 32: em_hum.*
- 33: em_hum.*
- 34: em_hum.*
- 35: em_hum.*
- 36: em_hum.*
- 37: em_hum.*
- 38: em_hum.*
- 39: em_hum.*
- 40: em_hum.*
- 41: em_hum.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	2310	100.0	2886	6	E58346	E58346 Nucleic aci
2	2310	100.0	2886	6	AX009610	AX009610 Sequence
3	2187.4	94.7	2907	3	AF272778	AF272778 Drosophil
4	1994.8	86.4	2834	3	AY036613	AY036613 Drosophil
5	776.2	33.6	42079	2	AC017620	AC017620 Drosophil
6	776.2	33.6	158758	3	AC092243	AC092243 Drosophil
7	776.2	33.6	186803	3	AC092225	AC092225 Drosophil
8	776.2	33.6	272521	3	AE003642	AE003642 Drosophil
9	776.2	33.6	320754	3	DROSADH02	AE003408 Drosophil
10	515.4	22.3	1683	3	DME554210	AE003408 Drosophil
11	512.8	22.2	3629	3	AF143846	AF143846 Heliothis
12	512.8	22.2	3700	6	AX009612	AF143846 Heliothis
13	512.8	22.2	3701	6	E58347	E58347 Nucleic aci
14	448	19.4	2023	3	AF321445	AF321445 Drosophil
15	441.6	19.1	2023	3	AF321447	AF321447 Drosophil
16	436.8	18.9	2023	3	AF321446	AF321446 Drosophil
17	433.6	18.8	1665	3	DME554209	AF321446 Drosophil
18	433.6	18.8	1699	3	BT011147	BT011147 Drosophil
19	424.6	18.4	2110	3	AF321449	AF321449 Drosophil
20	411.4	17.8	3029	3	AF143847	AF143847 Heliothis
21	411.4	17.8	3109	6	E58348	E58348 Nucleic aci
22	411.4	17.8	3109	6	AX009614	AX009614 Sequence
23	389.8	16.9	2068	3	AF321448	AF321448 Drosophil
24	321	13.9	1461	5	AY299465	AY299465 Takifugu
25	320.2	13.9	2116	3	AY036614	AY036614 Drosophil
26	316.4	13.7	1542	5	AY299466	AY299466 Takifugu
27	314	13.6	1719	5	AY247962	AY247962 Danio rer
28	313.8	13.6	2106	10	RATNARAD	L31619 Rattus ratt
29	313.8	13.6	3030	10	S53987	S53987 nicotinic r
30	311.4	13.5	1668	5	AY298752	AY298752 Takifugu
31	307.4	13.3	1848	10	MUSNARS	L37663 Mus musc
32	305.8	13.2	2088	10	AF225980	AF225980 Mus muscu
33	297.8	12.9	1559	9	HUMA7NAR	L25827 Human a7 ni
34	297.8	12.9	1916	9	AF385585	AF385585 Homo sapi
35	297.8	12.9	1977	9	HS040583	U40583 Human alpha
36	297.8	12.9	2087	9	HSARA7A	X70297 H.sapiens m
37	296.2	12.8	1509	6	AX054567	AX054567 Sequence
38	296.2	12.8	1509	9	HSNACHRA7	Y08420 H.sapiens m
39	296.2	12.8	1555	9	HSCHRNA7A	Z23141 H.sapiens C
40	296.2	12.8	1665	5	AY298753	AY298753 Takifugu
41	296.2	12.8	1876	6	AR055255	AR055255 Sequence
42	296.2	12.8	1876	6	AR071403	AR071403 Sequence
43	296.2	12.8	1876	6	AR173187	AR173187 Sequence
44	296.2	12.8	1876	6	AR224030	AR224030 Sequence
45	296.2	12.8	1876	6	AR282833	AR282833 Sequence

ALIGNMENTS

RESULT 1
E58346
LOCUS E58346
DEFINITION Nucleic acid encoding insect actyl choline receptor subunit.
ACCESSION E58346
VERSION E58346.1 GI:13019345
KEYWORDS JP 2000023680-A/1.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2886)
AUTHORS Martin,A., Nadja,E. and Thomas,S.
TITLE Nucleic acid encoding insect actyl choline receptor subunit


```
Db 2172 GGACGACGCTGATCTCTAGAGTCCGACACGCGCCCTGTTGCGACACATCTCCGACGG 2231
Qy 1861 AAGCACCAGATACCTCTCCGACGTTGAGCTGAAGAGCGCTCGTGAATCGCTCTGGCC 1920
Db 2232 AAGCACCAGATACCTCTCCGACGTTGAGCTGAAGAGCGCTCGTGAATCGCTCTGGCC 2291
Qy 1921 AACGTACTAGACATCGATGATGACTTCCGGACAAATTTGTCGCCCATGAGCCCGCGGA 1980
Db 2292 AACGTACTAGACATCGATGATGACTTCCGGACAAATTTGTCGCCCATGAGCCCGCGGA 2351
Qy 1981 ACATGCGCACACACCCCGCTTCTATCGACGCTTTATGACAGAGCGAGCATGGCAGC 2040
Db 2352 ACATGCGCACACACCCCGCTTCTATCGACGCTTTATGACAGAGCGAGCATGGCAGC 2411
Qy 2041 ATTGGGCAATTTGCGACACCCGCAATCGCGATCGCGTACCCATCATAGCTGCATCAA 2100
Db 2412 ATTGGGCAATTTGCGACACCCGCAATCGCGATCGCGTACCCATCATAGCTGCATCAA 2471
Qy 2101 TCATCACTGAATGATGATGAGTTTAACTCTTAAGGAAATTCGTTTATAACTGATCAG 2160
Db 2472 TCATCACTGAATGATGAGTTTAACTCTTAAGGAAATTCGTTTATAACTGATCAG 2531
Qy 2161 CTAGTAAAGATGACGAGTGAATGCAATGCGCAATGATGGAAATTTGCAAGCTATGGTC 2220
Db 2532 CTAGTAAAGATGACGAGTGAATGCAATGCGCAATGATGGAAATTTGCAAGCTATGGTC 2591
Qy 2221 GTTGACAGATGTCCTTATCATATTCACAAATGTCGAATATTAGCCACAAATAGCTGTA 2280
Db 2592 GTTGACAGATGTCCTTATCATATTCACAAATGTCGAATATTAGCCACAAATAGCTGTA 2651
Qy 2281 CTACTATCGGCACACATATTATGTCG 2310
Db 2652 CTACTATCGGCACACATATTATGTCG 2681
```

```
RESULT 2
AX009610
LOCUS AX009610 2886 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 1 from Patent EP0962528.
ACCESSION AX009610
VERSION AX009610.1 GI:9996842
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1
AUTHORS Adamczewski M.D., Schulte T.D. and Oellers N.D.
TITLE Nucleic acids encoding acetylcholin-receptor subunits from insects
JOURNAL Patent: EP 0962528-A 1 08-DEC-1999;
BAYER AG (DE)
FEATURES
source 1. .2886
Location/Qualifiers
/organism="Drosophila melanogaster"
/mol_type="unassigned DNA"
/db_xref="taxon:7227"
372..2684
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07499.1"
/db_xref="GI:9996843"
/db_xref="FEMTREMEL: CAC07499"
/translaton="MKNAQLKTEVDDDELWLAVRLAHCSNFSSSSSTRITSSNQHR
NQQLTLQPSRLSKHSNIASEQHSQOEPAKDEVDVNHGNSDQOHLQOOLDS
NMLSPKTAARADBATTCOPTNIRLCARORLRKRKRKPATNETDICKQQL
SMPPKTKSTDTYSTPATTCPTATYMOCRASDNESPIPSHRDVSSTATFVWLH
VLQVLLSLQWQLHVQORSVLLFRRIAATFIYSIGSFAQKNSSSSSSSNS
DPSQTLLINGLHMSWIFLLYLNLISAKVLGHEKRLDHLDDLPYNTLPERVLNES
DLQLSFLGLTMQI IDYDEKNQLLVNWLKLEWDMNLRWNTSDYGVKDLRIPPHR
IWKPDVLMYNSADSGFDGTQTNVVRNNGSLYVPFGIFKSTCKIDITWFPFDQRC
EMKFGSMTYDGFQDLQLQDFTGGDISSVLNGEWELLGVFGKRNELIYNCCPEPYID
```

CDS

```
ITFAIIRRETLXYFFNLIIPCVLLIASMALLGFTLPPDSGEKSLGVTILLSLTVFLN
MVAEIEMATSDAVPLMIRIVFLCWLPMILRMSRPGRELIIEFFITPDCSDTSSEKHKQI
LSVEIKRSKSSLLANLIDDDFRHNCRPMTPGGTLPHNPAPFRTVYGGDGSIG
PIGSTMPDAVTHHTTCIKSSTEYELGLILKEIRFITDQLRKDDCECNDIANDWKEAAMV
VDRCLLIIFTMFAILATIALLSAPHIIVS"
ORIGIN
Query Match 100.0%; Score 2310; DB 6; Length 2886;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 2310; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 ATGAAAAATGCACAACTGAACTGACTGAAGTTGACGATGATGAGCTGTGCTGCAGTA 60
Db 372 ATGAAAAATGCACAACTGAACTGACTGAAGTTGACGATGATGAGCTGTGCTGCAGTA 431
Qy 61 AGATTAGCGCACTGCGACGAGCAACTTTAGCAGCAGTAGCAGCAACAAGAACCCAGCAGC 120
Db 432 AGATTAGCGCACTGCGACGAGCAACTTTAGCAGCAGTAGCAGCAACAAGAACCCAGCAGC 491
Qy 121 AACGAGAGGCACACACGAGCAACTCACAACACTGCAACACGAGGAGCTTAAAGTACAAACAC 180
Db 492 AACGAGAGGCACACACGAGCAACTCACAACACTGCAACACGAGGAGCTTAAAGTACAAACAC 551
Qy 181 CACAGCAACATTTGCAAGCGAGCAGCACAATAGCCAGCAACAGGAGCCAGCATCGAAGGAC 240
Db 552 CACAGCAACATTTGCAAGCGAGCAGCACAATAGCCAGCAACAGGAGCCAGCATCGAAGGAC 611
Qy 241 GAGGATGAGCAACACGAGTAGAAGCAATGACAGCAGCAGCATCTGCAACAGCTAGAC 300
Db 612 GAGGATGAGCAACACGAGTAGAAGCAATGACAGCAGCAGCATCTGCAACAGCTAGAC 671
Qy 301 AGCAGCAACATTTGTCGCCAAGACAGCCGCGACGACCACTGCTGCGGCGATGAGCA 360
Db 672 AGCAGCAACATTTGTCGCCAAGACAGCCGCGACGACCACTGCTGCGGCGATGAGCA 731
Qy 361 ACAACCCCAACCAACCAATAGACTGTGTGCGACGCAAGCGCAACCAATTTGCTGCG 420
Db 732 ACAACCCCAACCAACCAATAGACTGTGTGCGACGCAAGCGCAACCAATTTGCTGCG 791
Qy 421 CGACGAAAAAGAAACACGACCAACCCCAACGAAACAGATATCAAGAAACACAGCAACTT 480
Db 792 CGACGAAAAAGAAACACGACCAACCCCAACGAAACAGATATCAAGAAACACAGCAACTT 851
Qy 481 AGCATGCTGCTTCAAAAACGCGCAATTCACGACCACTTACGACACACGAGGAGCAACA 540
Db 852 AGCATGCTGCTTCAAAAACGCGCAATTCACGACCACTTACGACACACGAGGAGCAACA 911
Qy 541 ACCAGCTGTGCGACGACCACTTACATGCAATGTGCGACGCGACCAATGAGTTTCAATTT 600
Db 912 ACCAGCTGTGCGACGACCACTTACATGCAATGTGCGACGCGACCAATGAGTTTCAATTT 971
Qy 601 CCGATATCGAGACATGATAGATATCCACGCCACATTCGCGCTGGGTGTTGCAATGCTG 660
Db 972 CCGATATCGAGACATGATAGATATCCACGCCACATTCGCGCTGGGTGTTGCAATGCTG 1031
Qy 661 CAGTGTGCTGCTGCTGCGTGCACAGTGCACACTTCACTGCAACAGCGATCGGTGCTA 720
Db 1032 CAGTGTGCTGCTGCTGCGTGCACAGTGCACACTTCACTGCAACAGCGATCGGTGCTA 1091
Qy 721 CTGTTTCAAGAGGATGCGCAGCGACCACTGCGCTTCAATTTCTATTTAGGCGAGCTTTGCA 780
Db 1092 CTGTTTCAAGAGGATGCGCAGCGACCACTGCGCTTCAATTTCTATTTAGGCGAGCTTTGCA 1151
Qy 781 GCGCACTGAAAAATAGCAGCAGCAGTAGCAGCAGCAACAGCAGCAACAACAGCAGC 840
Db 1152 GCGCACTGAAAAATAGCAGCAGCAGTAGCAGCAGCAACAGCAGCAACAACAGCAGC 1211
Qy 841 ACCGAAATATTAACGCACTTAATAACACTCATGATATTTTATTGATATATTTGAAT 900
Db 1212 ACCGAAATATTAACGCACTTAATAACACTCATGATATTTTATTGATATATTTGAAT 1271
Qy 901 TTATCTCTCTAAAGTTTCTCTAGCAGGATATCATGAAAGAGACTGTTTACAGCATCTTTTG 960
```

Db 1272 TTATCTGCTAAAGTTTGCTAGCAGGATATCATGAAAGAGAGACTGTTACACGATCTTTTG 1331
QY 961 GATCCTTATATACACTAGAACGTCCTCGTTCTCAATGATCGGACCCGTTACAAATTAAGC 1020
Db 1332 GATCCTTATATACACTAGAACGTCCTCGTTCTCAATGATCGGACCCGTTACAAATTAAGC 1391
QY 1021 TTTGGTTTAACTTTAATGCAAAATATCGATGTGGACGAGAAAAATCAATTTGCTAGTCACT 1080
Db 1392 TTTGGTTTAACTTTAATGCAAAATATCGATGTGGACGAGAAAAATCAATTTGCTAGTCACT 1451
QY 1081 AATGTGTGGTTAAACTGGAGTGGAAACGACATGAATCTCCGTCGGAACACCTCCCACTAT 1140
Db 1452 AATGTGTGGTTAAACTGGAGTGGAAACGACATGAATCTCCGTCGGAACACCTCCCACTAT 1511
QY 1141 GSCGGAGTTAGGATCTCGGATACGGCGGATCGCATCTGGAACCGGACGTCGTGATG 1200
Db 1512 GSCGGAGTTAGGATCTCGGATACGGCGGATCGCATCTGGAACCGGACGTCGTGATG 1571
QY 1201 TACAAACAGTGGGATGAGGGATTGACCGGACCTACCAAGACGAACTGTGTGTGGGAAAC 1260
Db 1572 TACAAACAGTGGGATGAGGGATTGACCGGACCTACCAAGACGAACTGTGTGTGGGAAAC 1631
QY 1261 AACGGCTCGTGTCTATACGTTCCCGCGGGGATCTTCAAGTCGACGTGCAAGATCGACATC 1320
Db 1632 AACGGCTCGTGTCTATACGTTCCCGCGGGGATCTTCAAGTCGACGTGCAAGATCGACATC 1691
QY 1321 ACGTGTTCCTCCCTTCGATGACGAGCGGTGCGAGATGAAGTTCGGCAGTTGGACCTACGAC 1380
Db 1692 ACGTGTTCCTCCCTTCGATGACGAGCGGTGCGAGATGAAGTTCGGCAGTTGGACCTACGAC 1751
QY 1381 GGATTCACAGCTGGATTTACAATTACAAGATGAAACTGGCGGTGATATCAGCAGTTTACGTTG 1440
Db 1752 GGATTCACAGCTGGATTTACAATTACAAGATGAAACTGGCGGTGATATCAGCAGTTTACGTTG 1811
QY 1441 CTCACCGCGAGTGGGAACTACTGGGTGTGCCGGGAAACGTAACGATCTATTACAAAC 1500
Db 1812 CTCACCGCGAGTGGGAACTACTGGGTGTGCCGGGAAACGTAACGATCTATTACAAAC 1871
QY 1501 TGCTCCCGGACCTTATAGACATCACTTCGCCATCATCATCGCGGACGAGACACTG 1560
Db 1872 TGCTCCCGGACCTTATAGACATCACTTCGCCATCATCATCGCGGACGAGACACTG 1931
QY 1561 TACTATTTCTTCAACCTGATCATACCTTTGTGTAATGATTCCTCCATGCGCTTGTCTCGGA 1620
Db 1932 TACTATTTCTTCAACCTGATCATACCTTTGTGTAATGATTCCTCCATGCGCTTGTCTCGGA 1991
QY 1621 TTCACCTGCGCCAGATTCGGGTGAAAAATTTATCGTGGGTGTTACCATCTTGCTCTCG 1680
Db 1992 TTCACCTGCGCCAGATTCGGGTGAAAAATTTATCGTGGGTGTTACCATCTTGCTCTCG 2051
QY 1681 CTGACCGTGTCTCTGAATATGTTCCGAGACAATGCGCGCTACTTCCGATCGCGTGCCA 1740
Db 2052 CTGACCGTGTCTCTGAATATGTTCCGAGACAATGCGCGCTACTTCCGATCGCGTGCCA 2111
QY 1741 TTGTGGATACGATCGTGTGTTTGTGCTGGTGCATGATATTCGAAATGAGTCCGCCCA 1800
Db 2112 TTGTGGATACGATCGTGTGTTTGTGCTGGTGCATGATATTCGAAATGAGTCCGCCCA 2171
QY 1801 GGACCAACCGCTGATCTAGAGTTCCGACCGACCGCCTGTTCCGACACATCTCCAGCGG 1860
Db 2172 GGACCAACCGCTGATCTAGAGTTCCGACCGACCGCCTGTTCCGACACATCTCCAGCGG 2231
QY 1861 AAGCAACCGATACTCTCCGACCTTGGAGTGAAGAGCGCTCGTCCGAAATCGCTGCGGC 1920
Db 2232 AAGCAACCGATACTCTCCGACCTTGGAGTGAAGAGCGCTCGTCCGAAATCGCTGCGGC 2291
QY 1921 AACGTACTAGACATCGATGATGACTTCCGGACAAATTTGCGCCCATGACGCCCGCGGA 1980
Db 2292 AACGTACTAGACATCGATGATGACTTCCGGACAAATTTGCGCCCATGACGCCCGCGGA 2351
QY 1981 ACACGTGCCACAACCCCGCTTCTATCGACGGTTTATGACAGGGACGATGGCAGC 2040
Db 2352 ACACGTGCCACAACCCCGCTTCTATCGACGGTTTATGACAGGGACGATGGCAGC 2411

QY 2041 ATTGGGCCAAATTGGCAGCACCCGAATGCGGATGCGGTCAACCATCATACGTGCATCAA 2100
Db 2412 ATTGGGCCAAATTGGCAGCACCCGAATGCGGATGCGGTCAACCATCATACGTGCATCAA 2471
QY 2101 TCATCAACTGAATATGAATTAGTTTAATCTTTAAAGGAAATTCGCTTTTATACTGATCAG 2160
Db 2472 TCATCAACTGAATATGAATTAGTTTAATCTTTAAAGGAAATTCGCTTTTATACTGATCAG 2531
QY 2161 CTACGTAAGATGACGAGTGAATGACATGTCGCAATGATGGAAATTTGCGAGCTATGGTC 2220
Db 2532 CTACGTAAGATGACGAGTGAATGACATGTCGCAATGATGGAAATTTGCGAGCTATGGTC 2591
QY 2221 GTTGACAGACTGTGCTTATCATATTCACATTTGCAATTTGCAATTTAGCCACCAATAGCTGTA 2280
Db 2592 GTTGACAGACTGTGCTTATCATATTCACATTTGCAATTTAGCCACCAATAGCTGTA 2651
QY 2281 CTACTATCGGCACCAATATTTATTTCTCG 2310
Db 2652 CTACTATCGGCACCAATATTTATTTCTCG 2681

RESULT 3

AF272778

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

1. 2907

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/chromosome="2"

/map="2L; 34E4-5"

/dev_stage="embryo"

1. 2907

/gene="nAcRalpha-34E"

/allele="A"

289. .2712

/gene="nAcRalpha-34E"

/note="ion channel; neurotransmitter transmembrane receptor"

/codon_start=1

/product="nicotinic acetylcholine receptor Dalphas subunit"

/protein_id="AAI13390.1"

/db_xref="GI:20152840"

/translation="MKNAQLKLTVEVDDDELMLAVLHCHSNFSSSSSTRTTSSNQRH NQQLITLQPSRLSTKHENLASEQHSQQSPASKDEVDVANHGRSNDQDTHLQDLSS NMLSPKTAATAAGDEATQOPTNRLICARRKRLRRRRKKRPATPNETDKKQOQL SMPFPKTRKSTDTYSTPAATTSCTPTATYMQCRASDNEFISIPISHRDVRVSTATFAWLH

AF272778 2907 bp mRNA linear INV 29-APR-2002
Drosophila melanogaster nicotinic acetylcholine receptor Dalphas
subunit (nAcRalpha-34E) mRNA, nAcRalpha-34E-A allele, complete cds.
AF272778
AF272778.1 GI:20152839

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 2907)

Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,

Dalphas, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a

New and Highly Conserved Target of Adenosine Deaminase Acting on

RNA-Mediated A-to-I Pre-mRNA Editing

Genetics 160 (4), 1519-1533 (2002)

21969411

11973307

2 (bases 1 to 2907)

Grauso, M. and Sattelle, D.B.

Direct Submission

Submitted (27-MAY-2000) Human Anatomy and Genetics, MRC-FGU,

University of Oxford, South Park Road, Oxford OX1 3QX, UK

Location/Qualifiers

1. 2907

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/chromosome="2"

/map="2L; 34E4-5"

/dev_stage="embryo"

1. 2907

/gene="nAcRalpha-34E"

/allele="A"

289. .2712

/gene="nAcRalpha-34E"

/note="ion channel; neurotransmitter transmembrane receptor"

/codon_start=1

/product="nicotinic acetylcholine receptor Dalphas subunit"

/protein_id="AAI13390.1"

/db_xref="GI:20152840"

/translation="MKNAQLKLTVEVDDDELMLAVLHCHSNFSSSSSTRTTSSNQRH

NQQLITLQPSRLSTKHENLASEQHSQQSPASKDEVDVANHGRSNDQDTHLQDLSS

NMLSPKTAATAAGDEATQOPTNRLICARRKRLRRRRKKRPATPNETDKKQOQL

SMPFPKTRKSTDTYSTPAATTSCTPTATYMQCRASDNEFISIPISHRDVRVSTATFAWLH

variation	variation	variation	variation	variation	variation	ORIGIN
<p>VLQVLVLSLQWQLHVQORSVLLFRRIAATSIAPISYLGSAFAQLKNSSSSSSSNS</p> <p>NPSTQLINGLNKHSITFLRLIYNLSAKVCLAGVHEKRLHLLDLPNTLTERPVLNES</p> <p>DEPLQSGFLTLMQIIVDVENKQLIVTNWLKLEWDMNLRWNTSDYGVKDLRIPIHR</p> <p>IKWDPVLWYNSADBGFGTQTNVVRNNGSVLPVPGI FKSTCKIDIWFPDFDORC</p> <p>EMKFGSWTQDFOLDLOLQETGGDISSVYLNGEWELLGYPGKRELIYNCCEPYID</p> <p>ITFAIIRRTLYYFNNLIIPCVLIASMAILGTLTPDPSGEKLSGLVITLLSUTFLN</p> <p>WVMTAETASDAVPLGLITFCNIMFWVASSVSKHLLNHYHNRNADTHESEWIRIV</p> <p>LCMLPWLIRMSRPGTILBFPTPCSDTSERKHQLSDVELKRSKSLLANVLDI</p> <p>DDDFHRNCRPMTFGTILPHNPAPFYRTVYGQDDGSGIPGSTRMPDAVTHHTCIKST</p> <p>EYELGLLKREITFRTDQLRKDDCEDNDIANDWKEAAMVVDRLCLIIFTMPAILLATI</p> <p>LSAPHIIVS"</p>	<p>373</p> <p>/gene="nAcRalpha-34E"</p> <p>/note="results in phenylalanine to isoleucine substitution; compared to B allele"</p> <p>375</p> <p>/replace="a"</p>	<p>375</p> <p>/gene="nAcRalpha-34E"</p> <p>/note="results in phenylalanine to isoleucine substitution; compared to B allele"</p> <p>827</p> <p>/replace="c"</p>	<p>827</p> <p>/gene="nAcRalpha-34E"</p> <p>/note="results in threonine to isoleucine substitution; compared to B allele"</p> <p>/replace="c"</p> <p>1079..1080</p> <p>/gene="nAcRalpha-34E"</p> <p>/note="results in lysine to arginine substitution; compared to B allele"</p> <p>/replace="gg"</p> <p>1081..1086</p> <p>/gene="nAcRalpha-34E"</p> <p>/note="results in deletion of asparagine and serine; compared to B allele"</p> <p>/replace="n"</p> <p>1119^1120</p> <p>/gene="nAcRalpha-34E"</p> <p>/note="results in serine, glycine, and asparagine insertion; compared to B allele"</p> <p>/replace="agcggcaac"</p> <p>1121</p> <p>/gene="nAcRalpha-34E"</p> <p>/note="results in asparagine to serine substitution; compared to B allele"</p> <p>/replace="g"</p>	<p>1079..1080</p> <p>/gene="nAcRalpha-34E"</p> <p>/note="results in lysine to arginine substitution; compared to B allele"</p> <p>/replace="gg"</p> <p>1081..1086</p> <p>/gene="nAcRalpha-34E"</p> <p>/note="results in deletion of asparagine and serine; compared to B allele"</p> <p>/replace="n"</p> <p>1119^1120</p> <p>/gene="nAcRalpha-34E"</p> <p>/note="results in serine, glycine, and asparagine insertion; compared to B allele"</p> <p>/replace="agcggcaac"</p> <p>1121</p> <p>/gene="nAcRalpha-34E"</p> <p>/note="results in asparagine to serine substitution; compared to B allele"</p> <p>/replace="g"</p>	<p>1121</p> <p>/gene="nAcRalpha-34E"</p> <p>/note="results in asparagine to serine substitution; compared to B allele"</p> <p>/replace="g"</p>	<p>Query Match 94.7%; Score 2187.4; DB 3; Length 2907;</p> <p>Best Local Similarity 95.4%; Pred. No. 0;</p> <p>Matches 2309; Conservative 0; Mismatches 1; Indels 111; Gaps 1;</p>
Qy	1	ATGAAATAATGCACAACTGCACTGAAAGTTTGCAGTATGATGAGCTGTGGCTGGCAGTA	60			
Db	289	ATGAAATAATGCACAACTGCACTGAAAGTTTGCAGTATGATGAGCTGTGGCTGGCAGTA	348			
Qy	61	AGATTAGCGCATGTCAGCAGCAACTTTAGCAGCAGTAGCAGCAAGAACCACCGACG	120			
Db	349	AGATTAGCGCATGTCAGCAGCAACTTTAGCAGCAGTAGCAGCAAGAACCACCGACG	408			
Qy	121	AACGAGGACAAACGAGCACTCAACACTGCACTCAACGAGGCTTAGTACAAACAC	180			
Db	409	AACGAGGACAAACGAGCACTCAACACTGCACTCAACGAGGCTTAGTACAAACAC	468			
Qy	181	CACAGCAACTTGCACGAGCAGCACTATAGCCAGCAACGAGGCGCAGCTCGAAGGAC	240			
Db	469	CACAGCAACTTGCACGAGCAGCACTATAGCCAGCAACGAGGCGCAGCTCGAAGGAC	528			
Qy	241	GAGGATGTAGCCAAACCGGTAGAACCAATGACCAGCAGCAGCATCTGCAACAGCTAGAC	300			
Db	529	GAGGATGTAGCCAAACCGGTAGAACCAATGACCAGCAGCAGCATCTGCAACAGCTAGAC	588			
Qy	301	AGCAGCAACTGTGTTCGCCAAAGACGCGGAGCAGCACTGTCTCGCGCGATGAAGCA	360			


```
|||||
Db 2553 AGACTGTCCTTATCATATTACAAATGTCGAAATATAGCAATATAGCAATAGCTGTACTACTA 2612
|||||
QY 2287 TCGGCACACATATATTGTCCTCG 2310
|||||
Db 2613 TCAGCACACATATATTGTCCTCG 2636
|||||

RESULT 5
AC017620 42079 bp DNA linear HTG 10-DEC-1999
LOCUS Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
DEFINITION AC017620
ACCESSION AC017620
VERSION AC017620.1 GI:6554377
KEYWORDS HTG; HTGS_PHASE2.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 42079)
Adams, M. and Venter, J.C.
Direct Submission
Submitted (10-DEC-1999) Celera Genomics, 45 West Gude Drive,
Rockville, MD, USA
This sequence was identified as CDM:10211513 by the submitter.
For more information on this record e-mail to fly@celera.com.
* NOTE: This is a 'working draft' sequence.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
FEATURES
Location/Qualifiers
source
1. 42079
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

ORIGIN
Query Match 33.6%; Score 776.2; DB 2; Length 42079;
Best Local Similarity 89.0%; Pred. No. 2.6e-201;
Matches 902; Conservative 0; Mismatches 13; Indels 98; Gaps 2;

QY 1 ATGAAAAATGCAACAACTGAACTGAAGTTGACGATGATGAGCTGTGGCTGGCAGTA 60
Db 10132 ATGAAAAATGCAACAACTGAACTGAAGTTGACGATGATGAGCTGTGGCTGGCAGTA 10191

QY 61 AGATTAGCGCACTGAGCGAGCACTTTAGCAGCAGTAGCAGCAACAGACCAACAGCAGC 120
Db 10192 AGATTAGCGCACTGAGCGAGCACTTTAGCAGCAGTAGCAGCAACAGACCAACAGCAGC 10251

QY 121 AACCAGAGGCAACACAGCAACTCAACACTGCAACAGGAGCTTAAGTACAAAACAC 180
Db 10252 AACAGAGGCAACACAGCAACTCAACACTGCAACAGGAGCTTAAGTACAAAACAC 10311

QY 181 CACAGCAACATTTGAAAGCGAGCAGCAATATGCCAGCAACAGGAGCCAGCATCGAAGGAC 240
Db 10312 CACAGCAACATTTGAAAGCGAGCAGCAATATGCCAGCAACAGGAGCCAGCATCGAAGGAC 10371

QY 241 GAGGATGTAGCAACACAGGTAGAGCAATGACCGAGCAGCAGCATCTGCAACAGCTAGAC 300
Db 10372 GAGGATGTAGCAACACAGGTAGAGCAATGACCGAGCAGCAGCATCTGCAACAGCTAGAC 10431

QY 301 AGCAGCAACATGTTGTGCGCAAGACAGCGCAGCAGCAACTGCTCCGGCGATGAAGCA 360
Db 10432 AGCAGCAACATGTTGTGCGCAAGACAGCGCAGCAGCAACTGCTCCGGCGATGAAGCA 10491

QY 361 ACAACCCACAAACCAACCAATAGACTGTGTGCAACGCAACAGATTCGCTCGC 420
Db 10492 ACAACCCACAAACCAACCAATAGACTGTGTGCAACGCAACAGATTCGCTCGC 10551

QY 421 CGACGAAAGAAACACGACCCCAACAGCAACAGCAATATCA-----463
Db 10552 CGACGAAAGAAACACGACCCCAACAGCAACAGCAATATCAAGGTAACATTAATAC 10611
```

```
QY 464 -----463
Db 10612 ATCATATTAACTAAATATAGAAATTTAGAAAATAATTGCACCTTCAGCACTTGAATT 10671

QY 464 -----AGAAACAAAGCAAGCACTTAGCATGCTCCCTTCAAAACGCGCA 505
Db 10672 TGGTCTTCTTACAAATTGCAGAAACAAACAGCAACTTAGCATGCTCCCTTCAAAACGCGCA 10731

QY 506 AATCCAGGACCACTTACAGCACCAGCAGCAACCAACAGCTGTCGAGCAGCACTTAC 565
Db 10732 AATCCAGGACCACTTACAGCACCAGCAGCAACCAACAGCTGTCGAGCAGCACTTAC 10791

QY 566 TGCATGTGCGAGCCAGCCAGCAATAGTTCAGTATTCGGATATCGAGACATGATAGATAT 625
Db 10792 TGCATGTGCGAGCCAGCCAGCAATAGTTCAGTATTCGGATATCGAGACATGATAGATAT 10851

QY 626 CCAGCGCCACATTCCTCCCTGGGTGTCATGTGCTCAGGTCGTGCTGCTGCTGCAAC 685
Db 10852 CCAGCGCCACATTCCTCCCTGGGTGTCATGTGCTCAGGTCGTGCTGCTGCTGCAAC 10911

QY 686 AGTGGCAACTTCACGTGCAACAGCGATCGGTCTACTGTTCAAGAGGATCGCAGCGAGCA 745
Db 10912 AGTGGCAACTTCACGTGCAACAGCGATCGGTCTACTGTTCAAGAGGATCGCAGCGAGCA 10971

QY 746 CCATCGCCTTCATTTCTTATTAGCAGCTTTGACGCGCAACTGAAAATAGCAGC---A 802
Db 10972 CCATCGCCTTCATTTCTTATTAGCAGCTTTGACGCGCAACTGAGGAGCAGCAGCAGTA 11031

QY 803 GCAGCAGTAGCAGCAGCAACAGCAGCAACAGCAGCAGCAACAGCAATATTAACGACATTA 862
Db 11032 GCAGCAGCAACAGCAGCAACAGCAGCGCGCAACAGCAGCAGCAACAGCAATATTAACGACATTA 11091

QY 863 ATAAACACTCATGGATATTTTATTGATATATTTTCAATTTATCTGCTCAAGTT 915
Db 11092 ATAAACACTCATGGATATTTTATTGATATATTTTCAATTTATCTGCTCAAGTT 11144
```

```
RESULT 6
AC092243/c
LOCUS Drosophila melanogaster, chromosome 2L, region 34D-34E, BAC clone
DEFINITION BAC09N24, complete sequence.
ACCESSION AC092243
VERSION AC092243.1 GI:14578124
KEYWORDS HTG
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 158758)
Celniker, S.E., Adams, M.D., Kronmiller, B., Tyler, D., Wan, K.H.,
Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C.,
Rogers, Y., An, H., Baldwin, D., Banzon, J., Beeson, K.Y., Busam, D.A.,
Carlson, J., Center, A., Champs, M., Davenport, L.B., Dietz, S.M.,
Dodson, K., Dorsett, V., Dou, L.E., Doyle, C., Dresnek, D., Farfan, D.,
Ferreira, S., Frise, E., Galle, R.F., Garg, N.S., George, R.A.,
Gonzalez, M., Houch, J., Hoskins, R.A., Hostin, D., Howland, T.J.,
Ibengwa, C., Jalali, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A.,
McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J.,
Pachleb, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B.,
Phouanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F.,
Stapleton, M., Strong, R., Svirska, R., Tector, C., Williams, S.M.,
Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.
Sequencing of Drosophila chromosome 2L, region 34D-34E
Unpublished
2 (bases 1 to 158758)
Celniker, S.E., Adams, M.D., Kronmiller, B., Tyler, D., Wan, K.H.,
Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C.,
Rogers, Y., An, H., Baldwin, D., Banzon, J., Beeson, K.Y., Busam, D.A.,
Carlson, J.W., Center, A., Champs, M., Davenport, L.B., Dietz, S.M.,
Dodson, K., Dorsett, V., Dou, L.E., Doyle, C., Dresnek, D., Farfan, D.,
```

Ferriera, S., Friese, E., Galle, R.F., Garg, N.S., George, R.A., Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J., Ibgwegam, C., Jalali, M., Kruse, D., Li, P., Mattai, B., Moshrefi, A., McIntosh, T. C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J., Pacleb, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B., Phoumanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F., Stapleton, M., Strong, R., Svirskaia, R., Tector, C., Williams, S.M., Zaveri, J.S., Smith, H.O., Rubin, G.M. and Venter, J.C.

Submitted (30-JUN-2001) Berkeley Drosophila Genome Project, MS 64-121, Lawrence Berkeley National Laboratory, One Cyclotron Road, Berkeley, CA 94720, US
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121

This sequence was assembled using end sequences from a whole genome shotgun and from subclones of this BAC and its neighboring clones. For further information about this sequence, including its location and relationship to other sequences, please visit our sequence archive Web site (<http://www.fruitfly.org/sequence/>) or send email to bdg@fruitfly.berkeley.edu.

FEATURES	SOURCE
----------	--------

```
1. 158758
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
/chromosome="2L"
/map="34D-34E"
/clone="R3C9N24 (D1347)"
/clone_lib="RPGI-98 (Roswell Park C
BR3C93.6)"
Drosophila melanogaster BAC library
```

ORIGIN

Query Match	33.6%;	Score 776.2;	DB 3;	Length 158758;
Best Local Similarity	89.0%;	Pred. No. 3.4e-201;		
Matches 902;	Conservative 0;	Mismatches 13;	Indels 98;	Gaps 2;
QY	1	ATCAAAATGCAACTGAACTGAACTGACTGAAGTTGACGATGAGCTGTGGCTGGCAGTA	60	
Db	14830	ATGAAAAATGCAAACTGAACTGAACTGACTGAAGTTGACGATGAGCTGTGGCTGGCAGTA	14771	
QY	61	AGATTAGCGCACTGCGAGCAGCAACTTTAGCAGCAGTACGACACAAAGAACACCAAGCAGC	120	
Db	14770	AGATTAGCGCACTGCGAGCAGCAACTCAGCAGCAGTACGACACAAAGAACACCAAGCAGC	14711	
QY	121	AACGAGGCGCAACACGAGCAACTCAACACTGCAACCAAGCAGCTTAAGTACAAAACAC	180	
Db	14710	AAACAAGAGGCGCAACGAGCAACTCAACACTGCAACCAAGGAGCTTAAGTACAAAACAC	14651	
QY	181	CACAGCAACATTGCAAGCGAGCGACACAAATAGCCAGCAACGAGGAGCCAGATCGAAGAC	240	
Db	14650	CACAGCAACATTGCAAGCGAGCGACACAAATAGCCAGCAACGAGGAGCCAGATCGAAGAC	14591	
QY	241	GAGGATGTAGCCAAACCGGTAGAGCAATGACAGCAGACGATCTGCCACACGCTAGAC	300	
Db	14590	GAGGATGTAGCCAAACCGGTAGAGCAATGACAGCAGACGATCTGCCACACGCTAGAC	14531	
QY	301	AGCAGCAACATTGTTGCGCCAAAGACAGCCGCGAGCAGCAACTGCTGCCGCGGATGAAGCA	360	
Db	14530	AGCAGCAACATTGTTGCGCCAAAGACAGCCGCGAGCAGCAACTGCTGCCGCGGATGAAGCA	14471	
QY	361	ACAAACCCCAACCAACCAACATAGACTGTGTGCACGCAAGCAGCAACGATTCGCTCGC	420	
Db	14470	ACAAACCCCAACCAACCAACATAGACTGTGTGCACGCAAGCAGCAACGATTCGCTCGC	14411	
QY	421	CGACGAAAAAGAAAAACCAAGCAACCCCAACGAAACAGATATCA-----	463	
Db	14410	CGACGAAAAAGAAAAACCAAGCAACCCCAACGAAACAGATATCAAGGTTAAACATTTAATAAC	14351	
QY	464	-----	463	

Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K., Nusser, D.R., Pacle, J.M., Palazzolo, M., Pittman, G.S., Pan, S., Pollard, J., Puri, V., Reese, M.G., Reinert, K., Remington, K., Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kiamos, I., Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spradling, A.C., Stapleton, M., Strong, R., Sun, E., Swirskas, R., Tector, C., Turner, R., Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wasserman, D.A., Weinstein, G.M., Weissenbach, J., Williams, S.M., Woodgett, W., Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yeh, R.F., Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H., Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O., Gibbs, R.A., Myers, E.W., Rubin, G.M. and Venter, J.C.
 The genome sequence of *Drosophila melanogaster*
 Science 287 (5461), 2185-2195 (2000)
 20196006
 10731132
 2 (bases 1 to 272521)
 Celniker, S.E., Adams, M.D., Kronmiller, B., Wan, K.H., Holt, R.A., Evans, C.A., Gocayne, J.D., Amanatides, P.G., Brandon, R.C., Rogers, Y., Banzon, J., An, H., Baldwin, D., Banton, J., Beeson, K.Y., Busam, D.A., Carlson, J.W., Center, A., Champs, M., Davenport, L.B., Dietz, S.M., Dodson, K., Dorsett, V., Dou, L.E., Doyle, C., Dresnek, D., Farfan, D., Ferreira, S., Frise, E., Gallie, R.F., Garg, N.S., George, R.A., Gonzalez, M., Houck, J., Hoskins, R.A., Hostin, D., Howland, T.J., Ibegwam, C., Jalali, M., Kruse, D., Li, P., Mattei, B., Moshrefi, A., McIntosh, T.C., Moy, M., Murphy, B., Nelson, C., Nelson, K.A., Nunoo, J., Pacle, J., Paragas, V., Park, S., Patel, S., Pfeiffer, B., Phouanavong, S., Pittman, G.S., Puri, V., Richards, S., Scheeler, F., Stapleton, M., Strong, R., Swirskas, R., Tector, C., Tyler, D., Williams, S.M., Zaveri, J.S., Smith, H.O., Venter, J.C. and Rubin, G.M.
 Sequencing of *Drosophila melanogaster* genome
 Unpublished
 3 (bases 1 to 272521)
 Misra, S., Crosby, M.A., Matthews, B.B., Bayraktaroglu, L., Campbell, K., Hradscky, P., Huang, Y., Kaminker, J.S., Prochownik, S.E., Smith, C.D., Tupy, J.L., Bergman, C.M., Berman, B.P., Carlson, J.W., Celniker, S.E., Clamp, M.E., Drysdale, R.A., Emmert, D., Frise, E., de Grey, A.D.N.J., Harris, N.L., Kronmiller, B., Marshall, B., Millburn, G.H., Richter, J., Russo, S., Searle, S.M.J., Smith, E., Shu, S., Smutniak, F., Whitfield, E.J., Ashburner, M., Gelbart, W.M., Rubin, G.M., Mungall, C.J. and Lewis, S.E.
 Annotation of *Drosophila melanogaster* genome
 Unpublished
 4 (bases 1 to 272521)
 Adams, M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.
 Direct Submission
 Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA
 5 (bases 1 to 272521)
 FlyBase
 Direct Submission
 Submitted (06-SEP-2002) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA
 6 (bases 1 to 272521)
 FlyBase
 Direct Submission
 Submitted (23-JAN-2003) University of California Berkeley, 539 Life Sciences Addition, Berkeley, CA 94720, USA
 On Sep 16, 2002 this sequence version replaced gi:7298121.
 Location/Qualifiers
 1. 272521
 /organism="Drosophila melanogaster"
 /mol_type="genomic DNA"
 /db_xref="taxon:7227"
 /chromosome="2L"
 /note="genotype: Y[1]; cn[1] bw[1] sp[1]; Rh6[1]"
 2745. 4311
 /locus tag="CG31765"
 /note="synonym: CG8855; last curated on Thu Nov 07 12:45:45 PST 2002"
 /map="34ES-34E5"

/db_xref="FLYBASE:FBgn0051765"
 join(2745. 2799,2858. 2975,3043. 3746,3914. 4311)
 /locus tag="CG31765"
 /product="CG31765-RA"
 /db_xref="FLYBASE:FBgn0051765"
 join(2745. 2799,2858. 2975,3043. 3746,3914. 4311)
 /locus tag="CG31765"
 /note="CG31765 gene product"
 /codon_start=1
 /product="CG31765-PA"
 /protein_id="AAN10860.1"
 /db_xref="GI:22946455"
 /db_xref="FLYBASE:FBgn0051765"
 /translation="MPDICSCEGYGKHCTQRCRDHWGLDCKNLCCQGAACDNL
 SGLCHTAGWTGQCELPQGTGIMCKKACDCDEKPCNPTGACIQDQPLQNLVS
 HVIETVNSTLKMGIIPRTTPVPLPEVIKQPTSNENAHGPKIIVHSSSLLLE
 LMLLAAAGVPTPEVIHVTNGITSPOEHLAGFYGGSEANSOTATADHSGLVLTLS
 LPEPVAFVTPQSENAQPELYDSFNNSSIKTPYAYARKESLYSVVSKRKGHEP
 SHLDELRIHQOHHQHPHQSHQHVQHSHTASQFLPARPQIMPANDQFSNGSVHPHQ
 RSHVTHLIIPPPQNSNLFQVPAQITAKRAHL"
 5340. 6882
 /gene="BG:DS00180.12"
 /locus tag="CG33119"
 /note="synonym: CG8855; last curated on Thu Nov 07 12:47:27 PST 2002"
 /map="34ES-34E5"
 /db_xref="FLYBASE:FBgn0027929"
 join(5340. 5532,5693. 6250,6329. 6532,6591. 6882)
 /gene="BG:DS00180.12"
 /locus tag="CG33119"
 /product="CG33119-RA"
 /db_xref="FLYBASE:FBgn0027929"
 join(5342. 5532,5693. 6250,6329. 6532,6591. 6846)
 /gene="BG:DS00180.12"
 /locus tag="CG33119"
 /note="BG:DS00180.12 gene product"
 /codon_start=1
 /product="CG33119-PA"
 /protein_id="AAF53360.2"
 /db_xref="GI:22946456"
 /db_xref="FLYBASE:FBgn0027929"
 /translation="MHCSGFIAGLTALLTLVAFVQIKDSTVSTELYGDIENYEYGD
 FESLSQREKQHLCHREVPFQFQREDFVRNGSTIYFHRIEVCAGYRDPYAN
 ECPVDSASPDNCRNGFCRSPGVCEFAEVRNEHGACHTPCIAQHGRCYLGTC
 VHQNFVLDQTRQFCRKGCSQSCGTCHECVAPGQCDSPGYRRTPLDGCOPVCPADC
 GFGKVPANQCECFAGIKRPNWVCEAEVLCNGLCRSRYKCHREGYRYDVNTT
 SCLPECSNDGCGNGVCIAPGVCRCFEGYVHGAECFKCEFCGKGRVCVAPICGC
 GEGQHCNGSCDDIEHCSPSGETHFIDRLKADRLSQHLNISEKRNHRQLAYEF
 NALIGLRFNF"
 complement(6800. 8816)
 /gene="BG:DS00180.7"
 /locus tag="CG31839"
 /note="synonym: CG16882; last curated on Mon Nov 04 13:31:36 PST 2002"
 /map="34ES-34E5"
 /db_xref="FLYBASE:FBgn0028543"
 complement(join(6800. 7190,7269. 7805,8300. 8816))
 /gene="BG:DS00180.7"
 /locus tag="CG31839"
 /product="CG31839-RA"
 /db_xref="FLYBASE:FBgn0028543"
 complement(join(6881. 7190,7269. 7805,8300. 8718))
 /gene="BG:DS00180.7"
 /locus tag="CG31839"
 /note="BG:DS00180.7 gene product"
 /codon_start=1
 /product="CG31839-PA"
 /protein_id="AAN10861.1"
 /db_xref="GI:22946457"
 /db_xref="FLYBASE:FBgn0028543"
 /translation="MBSSSCQLVTLGVLLAI CSLGGQPKTAGIKTRPPSGNLQJAG
 NSSESGRWSYNSGWSYQNSQNSYANQNVHVEQSGAGFVRAEVFPVTLPLPYGHY

COMMENT

On or before Mar 22, 2000 this sequence version replaced
gi:2337897, gi:1945578, gi:2337896, gi:2347073, gi:2337894,
gi:3097819, gi:895619, gi:3097823.
Submitted by the Berkeley Drosophila Genome Project. For more
information, visit the BDGP Web site: <http://www.fruitfly.org/> This
is the finished sequence of 3A4C-36A7.
The orientation of this sequence along the chromosome is left to
right. This sequence was annotated by Sima Misra
(sima@fruitfly.berkeley.edu) on behalf of the Berkeley Drosophila
Genome Center. Coding sequences are predicted based on
computational analysis, using both gene and CDS prediction programs
and matches to other sequences. These predictions and matches have
been evaluated by the annotators and may have been refined by hand.
The annotators have also used their judgement about which matches
to include in this record. The annotations about which matches
examined in more detail from
<http://www.fruitfly.org/publications/Adh.html>
The annotation syntax used in this record is documented at
ftp://ftp.ebi.ac.uk/pub/databases/edgp/sequence_annotation.README.v
1.2.

FEATURES
source

```
1. 320754
   Location/Qualifiers
     clone="P1 DS01368 (D100), P1 DS08249 (D25), P1 DS08284
       (D247), P1 DS00941 (D28), P1 DS08220 (D39), P1 DS00180
       (D29), P1 DS01514 (D33), P1 DS00131 (D35), P1 DS05899
       (D22), P1 DS01759 (D41), P1 DS01523 (D34), P1 DS01652
       (D43), P1 DS03792 (D31), P1 DS01068 (D37), P1 DS06238
       (D26), P1 DS08340 (D38), P1 DS04641 (D30), P1 DS01160
       (D42), P1 DS01486 (D32), P1 DS09219 (D101), P1 DS07721
       (D20), P1 DS00810 (D21), P1 DS06874 (D19), P1 DS03431
       (D102), P1 DS03144 (D106), P1 DS03323 (D127), P1 DS01219
       (D129), P1 DS00929 (D48), P1 DS04929 (D104), P1 DS03192
       (D103), P1 DS09194 (D267), P1 DS07295 (D47), P1 DS05639
       (D11), P1 DS07851 (D49), P1 DS03023 (D8), P1 DS01845
       (D7), P1 DS04862 (D107), P1 DS07108 (D51), P1 DS09217
       (D109), P1 DS02252 (D97), P1 DS00365 (D108), P1 DS07486
       (D46), P1 DS00913 (D24), P1 DS04095 (D112), P1 DS02795
       (D50), P1 DS07473 (D113), P1 DS02740 (D27), P1 DS09218
       (D114), P1 DS02780 (D23), BAC BACR48E02 (D482), BAC
       BACR44L22 (D483)"
     clone_lib="P1 library, partial Sau3a in pMS582tet14Ad10
       and RPl-38 (Roswell Park Cancer Institute Drosophila
       melanogaster BAC library, partial EcoRI in pBACe3.6)"
     complement(135..1376)
     gene="adat"
     join(135..447,505..1376)
     note="symbol=adat; synonym=BG:DS00941.2;
       prediction=method:genscan, version:1.0; Date run:
       19-Nov-98; Time: 15:48:39, score:54.110;
       match=method:tblastx, version:2.0al9MP-WashU
       [05-Feb-1998] [Build sol2.5-ultra 01:47:35 05-Feb-1998],
       score:329.0, desc:GenBank:U99383.H.sapiens mRNA for
       RNA-specific editase. CDS:148..2253; PID:e254628;
       PID:gi707504, species:Homo sapiens"
     codon_start=1
     product="hypoetical protein"
     protein_id="AAF44818.1"
     db_xref="GI:7287780"
     db_xref="FLYBASE:FBgn0028658"
     translation="MCDNKKPTVKEIAELCLKKFESLPTGKPTANOWTILAGVFN
       NNTSLQVSLGGTCKIGESKLCNGLILNDASHAVLARRGLRFLYQLKODRIFH
       RNTATSYMDGHEVEHFLSTQPCGACILEEQFAARAKRORDEDSMTYGAFL
       ISLSDPMLQTCALTKPGRGSHVSCSDKIARNVIGVQALLDVLISKPTIF
       SFLNFCDDAQLSELEAIFKDFCTFKHTRFPQRPQINIDPGRFFFSQSDWQP
       SPNGLIWSQVPEELRPEYISVNGKROQVTKKMKTSQAALAIKSKYLFTFLVLFVN
```

gene
CDS

gene
CDS

```
PLSEMFQQQLSDPERIAYASCKLDARYQFAWRBIKEXYFLQWTKPKHELLDFNPM5
NK"
1571..2560
  /gene="BG:DS00941.3"
  join(1571..1691,1752..1912,1978..2283,2405..2560)
  /note="symbol=BG:DS00941.3; prediction=method:genscan,
    version:1.0; Date run: 19-Nov-98; Time: 15:48:39,
    score:99.140, match=method:BLASTX,
    version:2.0al9MP-WashU [05-Feb-1998] [Build sol2.5-ultra
    01:47:30 05-Feb-1998], score:121.0,
    desc:trEMBL:g272565.MATCH TO EST AA085966 (FRAGMENT).
    organism:HOMO SAPIENS (HUMAN). dbxref:GenBank; AC004000;
    g272565; -, species:HOMO SAPIENS"
  /codon_start=1
  /product="hypoetical protein"
  /protein_id="AAF44819.1"
  /db_xref="GI:7287781"
  /db_xref="FLYBASE:FBgn9999001"
  /db_xref="FLYBASE:FBgn0028919"
  /translation="MPKVVSRSIVCSDTKQOEYNEEKPLNIYVCLNMKALILDCTL
    EQLPLREVDNARVINANDHANKLTHNPTPRMVYIKRSRNGIKKQYKRSCLPL
    YRHSPDSHVTVMNSALIRNKGESPLTOLLNSIKGSFKAPAKPATISAGPDSGIV
    DAGKKVVTHTKMGKFSVYSTVIDEEDIEAREIADSIANNARIKQLQKRG
    KLSLDVGIKTKIEDAPPPQKQKQRTLLER"
  join(3739..5231,5285..5642,5705..6627,6692..7540,
    7603..8493,8562..9195,9261..9751)
  /note="symbol=Sos"
  4140..9304
  /gene="Sos"
  join(4140..5231,5285..5642,5705..6627,6692..7540,
    7603..8493,8562..9195,9261..9304)
  /note="symbol=Sos; synonym=BG:DS00941.4;
    match=method:sim4, score:1000.0,
    desc:GenBank:M83931.Drosophila melanogaster son of
    sevenless (Sos) mRNA, complete cds. CDS:346..5133;
    PID:g158485, species:Drosophila melanogaster";
  match=method:BLASTX, version:2.0al9MP-WashU
  [05-Feb-1998] [Build sol2.5-ultra 01:47:30 05-Feb-1998],
  score:889.0, desc:trEMBL:007889.SON OF SEVENLESS 1
  PROTEIN (SOS 1), dbxref:GenBank; L13857; 9306778; -,
  species:HOMO SAPIENS"
  /evidence=experimental
  /protein_id="AAF44820.1"
  /db_xref="GI:7287782"
  /db_xref="FLYBASE:FBgn9999000"
  /db_xref="FLYBASE:FBgn0001965"
  /translation="MFGSPGSHAHTISYGGGIGLGTGGGGSGSGSGSGGGGIGI
    GGGVAGLQDCDGYDTKCNAAWRGLFTPSLLKVLQVHPRTAKEDALLYVKLIC
    LRLAMLCAPLPHSVQDVEEKVNSFPAPIDQWALNKAKEVINSKKRKSVPTEKVH
    TLQKDVLOYKIDSSVSAPLVALEVISADILKMGADYVIKIAHCETTKEDIEMVNA
    DVLMDMLNOSSEAHILPSPLPAQASATVEETVKELHDEKQYQCDLHWIRVFE
    EFDVYKDAPELEPIFSNIMDIYEVTVLLGSLEDVIEMSQSAPCVSCFELABA
    HARVFDYIKHLDSSQDDIESFEQQLGLHLHCDLEKVMASLSEKQVPSGRV
    RROLAERTRELQKQVHEDKQNCNEFREDISLKSGLSKGRKTSBERKVFLLFGL
    MVLCKANTKQTSAGATAYDLREKQYFMRVDINDRPSDDLKNSFELAPBMQPI
    VLTAKNAOHKHDMAILLAVITKSLDRLHLSLQDIERKHLRMPSPETIKFAVDS
    GDNVILBEESAGVPMIKATICKLERLTHYIADPTFVRTITTYIFPCSPQQLLQ
    LLVERFNIDPSLVQDTGTAGAGMGVGGDKHKNSHREDWKRYKQVQVQVRV
    LNVLRHWDHDFYDFEKDPMLEKLFNLHLEHNGKSMRWDSVLKIVQRKQEKSN
    KTIYAYGHDPPIEHLSPNDEITLLTHLPLEARQLTLEFEEKNVKNKPSLVS
    PTKDKVEKSNLLKIMKHTNTVTVEIKSITAEENYBERLAIMQRAIEVMVWLEL
    NNFNGILSIVAANGTASVRLRWTFGLPERYRKFLIECRELSDDLKIKERLSIN
    PCVPFFGRLTNILHLEGNPDLLANTEINFSRRKRVAEIIGETQQYQNYCINE
    ESTIQFFQLQDFNGLSDKSDYLYNESLRIEPRCKTVKPFPRKPHIPLKSPGI
    KPRQNTNSSKLSNSTSVAASAAASATSIATASAPLSHSAIMDAPAAAAA
    GSGTLAGESQPHNHAFSVAPVPIPERNTSSWSTPOHTRTDONNGVSVPAHLF
    KPGGAHWANNSTLASASAMDVPSPALPEHLPOSLPDSNPPASDTEAPSLPLK
    VVSPRHETGNRSPFHGRMQNSTHSTASTVTLTGSTSGEEFCAGGFIENSAHQSP
    GAVPISPHVNPAINMETRAYPPPLPPRRKERTESCADMAQKQAPDAPTPPRDGE
```



```
Db 545 ATCGAGTGGAGAGAGAACAGAGCTTTTATAAACAACATCTGGCTAAAACTAGAGTGG 604
QY 1105 AACGACATGAATCTCGCTGGAGAACCTCCGACTATGGCGAGTTAAGGATCTGGAATA 1164
Db 605 AATGATATGAATCTGAGGTGGAACTCTCAGATTTTCGGCGGGTCAAAAGATTAAAGAGTG 664
QY 1165 CGCGCGCATCGCATCTGGAAGCCGAGCTGCTGATGTACAAAGTGGCGGATGAGGGATT 1224
Db 665 CCACCCCAAGACTATGGAACCCAGAGCTCTTATGTATCAACAGCGCGGACGAAGGGTTC 724
QY 1225 GACGCACTTACCAGACGAAGTGGTGGTGGGAAACAACGGCTCGTGTCTATACGTTCCG 1284
Db 725 GACAGCAGTATCCAAACGAAGTGGTGGTGGGAAACAACGGCTCGTGTCTGTACGTCGG 784
QY 1285 CGCGGATCTTCAAGTCCAGCTGCAAGATGCACATCAGTGGTTCCCTTCGATGACCAAG 1344
Db 785 CCGCGCATCTTCAAGAGCACTGCAAGATGCACATCAGTGGTTCCCTTCGACGACCAA 844
QY 1345 CGGTGCGAGATGAAGTTCGGCAGTTGGACCTACGACGGATTCCAGCTGGATTACAAATTA 1404
Db 845 CGATGCGAGATGAAGTTTGGCAGCTGGACTTATGATGGTTATCAGTTGGATCTACAATA 904
QY 1405 CAAGATGAACCTGGCGTGATATGACAGTTACGTGCTCAACGGCGAGTGGAACTACTG 1464
Db 905 CAGGATGAAGGGGGGAGATATAAGCAGTTTGTACGAATGGCGAATGGAGTTAATA 964
QY 1465 GGTGTGCCCGCAACAGTAAACGAGATCTATTACACTCTGCCGGAACCTATATAGAC 1524
Db 965 GGAGTCCCGGAGCGACAGATCTACTACACTGTGTGGGAGCCATACATCGAC 1024
QY 1525 ATCACTTTCGCCATCATCCGCGGAGAACTGTACTATTCTTCAACCTGATCATA 1584
Db 1025 ATCACTTTCGGTGGTGGTATCCGAGGAGAAAGCTCTACTACTTCTTCAATCTGATG 1084
QY 1585 CTTGTGTACTGATGCTTCCATGCTGCTGCTGGATTCACCTGCGCGCCAGATTCGGGT 1644
Db 1085 CCGTGGCTGCTCATCGCTCCATGCTCTATTGGGGTTTCACTTGGCTCCAGACTCCGGA 1144
QY 1645 GAAAAATATTCGCTGGGTGTTTACCATCTGCTCTGCTGACCGTGTCTTCAATGGTT 1704
Db 1145 GAAAGTTGCTTTAGGTGTGACGATATCTGCTGTTGAGGGTGTCTTCAACATGGTG 1204
QY 1705 GCGAGACAATGCGCGTACTTCCGATCGGTGGCAATG-----TGATACGC 1752
Db 1205 GCGAGACGATGCGCAGCAGCTCGAGCGCGTGGCTTCTCGCAGCTACTTCACTGC 1264
QY 1753 ATCGTGTGTTTG--TGCTGGTGCATGATGATTTGGAATGAGTGCOCAGGACGACGC 1810
Db 1265 ATCATGTTTCATGGTGGTCTTCTCGTCTGCTCCACCATCTGATCTCTCACTACCA 1324
QY 1811 TGATCTCTAGAGTTCGCGACAGCGCTTTCGAGACATCTCCGAGGGAA-----GCAC 1866
Db 1325 CGGACGCGACACACTCAGAAATGAGTATGGAATCGTTGGCGGTTCCTTTATGGCTG 1384
QY 1867 CAGATCTCTCGAAGTGTGAGCTGAAAGAGCGCTCGTCAAAATCGCTGCTGGCGCAACGTA 1926
Db 1385 CCGTGGTGTGTCGATGTCAACGCGCGGCTCGCGAGAGCGCGCGCGCGCGCGCGTA 1444
QY 1927 CTAGACATCATGATGATCTCCGGCAAAATGTGCGCCCATGAGCGCCCGCGGAAACTG 1986
Db 1445 CTCCGCGCGCGGACTGAGCTGCGGAGCGCTCTCTCAAGTCTCTCTAGCGAAGCTG 1504
QY 1987 CCACAAACCCGGCTTCTATCGACGCTTTATGGAAGCGGACGATGGACGATGGG 2046
Db 1505 CTGACATCATGACGACTTCCGCGCACCGGAGCGGAGCGGCGCAATGCTCCGATAC 1564
QY 2047 CCAATGGGACGCCGGAATGCCGATGCGGTCAACCATCATACGTGTCATCAATCATCA 2106
Db 1565 TACAGGCGGGGTGAGGAGATGCGCGGGTGGCGCGCACAGTTGCTTC-----GGT 1618
QY 2107 ACTGAATATGAATTAGTTTAAATCTTAAAGAAATTCGCTTTTATACTGATCAGTACGT 2166
Db 1619 GTCGACTACGAGTCTCCCTCATCTCTGAGGAGATTAGATCATCAGATCAGATGCGC 1678
```

```
QY 2167 AAGATGACGATGCAATGATCCCAATGATTTGAAATTTGACGATATGTCGTGAC 2226
Db 1679 AAGGACGACGAGATGCGGACATTTCCGCGACTCGAAGTTCCGCCCATGTCGTGAC 1738
QY 2227 AGACTGTCGCTTATCATATTACATGTTTCCAAATTTAGCCCAATAGCTGTACTATA 2286
Db 1739 AGACTGTCGCTTATTTATCTTTACCGTGTTCACATCATCGCACGCTAGCCGTGCTGTG 1798
QY 2287 TCGGCACCATATATTATTGTCG 2310
Db 1799 TCGGCGCCACACATCATGTCG 1822

RESULT 13
E58347
LOCUS E58347 3701 bp DNA linear PAT 18-JUN-2001
DEFINITION Nucleic acid encoding insect actyl choline receptor subunit.
ACCESSION E58347
VERSION E58347.1 GI:13019346
KEYWORDS JP 2000023680-A/2.
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1 (bases 1 to 3701)
AUTHORS Martin,A., Nadja,E. and Thomas,S.
TITLE Nucleic acid encoding insect actyl choline receptor subunit
JOURNAL Patent: JP 2000023680-A 2 25-JAN-2000;
BAYER AG
COMMENT
OS Heliothis virescens
PN JP 2000023680-A/2
PD 25-JAN-2000
PF 26-APR-1999 JP 1999118159
PR 04-MAY-1998 DE 19819829.9
PI MARTIN ADAMUTSUEUSUKI,NADJA ERASU,THOMAS SCHULTE PC
C12N15/09,A01K67/033,C07K14/705,C07K16/28,C12N1/21,C12N5/10, PC
C12Q1/68,
PC G01N33/15,G01N33/50///(C12N1/21,C12R1/19),C12N15/00,C12N5/00 CC

FH Key Location/Qualifiers
FT CDS (335)..(1822).
FEATURES
source
1..3701
/organism="Heliothis virescens"
/mol_type="genomic DNA"
/db_xref="taxon:7102"
ORIGIN
Query Match 22.2%; Score 512.8; DB 6; Length 3701;
Best Local Similarity 62.5%; Pred. No. 4.3e-129;
Matches 878; Conservative 0; Mismatches 502; Indels 24; Gaps 4;
QY 925 GGATATCATGAAAAGAGACTGTGTACAGGATCTTTTGGATCCTTATATACACTAGAACGT 984
Db 425 GGGTACACGAGAGAGCGGCTACTGCACACTTATGGACCTACACGCTACTGGAGG 484
QY 985 CCGGTTCTCAATGAATCGGACCGTTTACAAATTAAGCTTTTGGTTTAACTTTAATGCAATT 1044
Db 485 CCGTGTCTCAACGAGAGCGACCGCTCTCTTGGCTCAGCTCATGCGATC 544
QY 1045 ATCGATGTGGACGAGAAAAATCAATTGCTAGTCACTTAATGTGTGGTTAAAACTGGAGTGG 1104
Db 545 ATCGAGTGGACGAGAGAACCGAGCTTTTAAATACAAACATCTGGCTAAACTAGAGTGG 604
QY 1105 AACGACATGAATCTCCGCTGGAAACCTCCGATATGGCGGATTAAGGATTCGGAATA 1164
Db 605 AATGATATGAATCTGAGGTGGAACACTTCAGATTTTCGGCGGGTCAAAAGATTAAAGAGTG 664
QY 1165 CGCGCGCATCGCATCTGGAAGCCGAGCTGCTGATGTACAAAGTGGCGGATGAGGGATT 1224
Db 665 CCACCCCAAGACTATGGAACCCAGAGCTCTTATGTATCAACAGCGCGGACGAAGGGTTC 724
```


/gene="nAcRalpha-30D"
/note="results in asparagine to serine substitution;
compared to variant clone"
/replace="g"
1314
/gene="nAcRalpha-30D"
/note="compared to variant clone"
/replace="a"
1560
/gene="nAcRalpha-30D"
/note="compared to variant clone"
/replace="t"
1605..1607
/gene="nAcRalpha-30D"
/note="results in serine deletion; compared to variant
clone"
/replace=""
1683
/gene="nAcRalpha-30D"
/note="compared to variant clone"
/replace="t"
ORIGIN

Query Match 19.4%; Score 448; DB 3; Length 2023;
Best Local Similarity 58.5%; Pred. No. 2.5e-111;
Matches 828; Conservative 0; Mismatches 570; Indels 18; Gaps 2;

Qy 909 TAAAGTTTGGCTACAGGATATCATGAAAAGAGACTGTTACAGCATCTTTGGATCCCTTA 968
Db |||||
Qy 441 TAAAGAAAGTGTCAAGGACCTCTGAAAGCGCTGCTGAACCATCTGCTGTGCCACCTA 500
Db |||||
Qy 969 TAATACACTAGAACGTCCTGTTCTCAATGAATCGGACCGTGTACAATTAAGCTTTGGTTT 1028
Db |||||
Qy 501 CAATACGCTGAGCGACCCGTGGCAATGAATCGAGCCCTCGAGGTTAAGTTGCGGACT 560
Db |||||
Qy 1029 AACTTTAATGCAATATPCGATGTGACGAGAAAATCAATTTGTAGTCACTAATGTGTG 1088
Db |||||
Qy 561 GACGCTGAGCAGATCATCGAGTGGATGAAAGAAATCAGATTTCTGACCACAAATGCGTG 620
Db |||||
Qy 1089 GTTAAACTGGAGTGGACGACATGATCTCCGTGGAACACCTCCGACTATGCGGAGT 1148
Db |||||
Qy 621 GTTAAATTTGGAGTGGAAACGACTCAATCTCGCTGGAATGAAACGGAATACGCGCGGT 680
Db |||||
Qy 1149 TAAGGATCTCGAATACCGCGCATCTGCAAGCCGACCTGCTGATGTGTACACAG 1208
Db |||||
Qy 681 CAAGGATCTAGATACGCCCACAGCTGTGGAAGCCCGACGTGCTCATGTACACAG 740
Db |||||
Qy 1209 TCGGATGAGGATTTGACGSCACTACAGACGAGCTGTGTGCGGAAACACGCGTC 1268
Db |||||
Qy 741 CGCGGATGAGGATTCGATGSCAGCTATCACCAACATTTGTGTGTCACAAATACAGCGAG 800
Db |||||
Qy 1269 GTGTCTAGTTCGCGCGGGATCTTCAAGTCGACGTGCAAGATCGACATCGAGTGGTT 1328
Db |||||
Qy 801 TTGTCTGTAGTGCCTGCTGATCTTCAAGAGACACATGCAAGATAGACATCAAGTGGTT 860
Db |||||
Qy 1329 CCCCTTCGATGACAGCGGTGGAGATGAAAGTTGCGGAGTTGGACCTACGACGATTCOA 1388
Db |||||
Qy 861 CCCATTGTATGACCAACATTTGGGAATGAAATTCGGTAGTTGACTTACGATGGAATCA 920
Db |||||
Qy 1389 GCTGGATTTCAATTAAGATGAAACTGGCGGTGATATCAGCAGTTACGTGTCTCAACGG 1448
Db |||||
Qy 921 GTTGGATTTGGTTTGAATTCGGAAGATGAGGGGATCTTCGATTTTCAACAATGG 980
Db |||||
Qy 1449 CGAGTGGAACTACTGGTGTGCCCGCAACGTAACGATCTATTACAACTGTGTCGCC 1508
Db |||||
Qy 981 CGAGTGTACTTGTCTGCCATGCGGGAAGAAATACGATAGTCTACGCTGTGTCGCC 1040
Db |||||
Qy 1509 GGAACCTTATAGACATCACTTCGCATCATCATCCGCGGACGACACTGTACTATTT 1568
Db |||||
Qy 1041 AGAACCATATGCGATATCACTTTACTATACAAATTCGTGCGCGTACATATATTT 1100
Db |||||
Qy 1569 CTTCAACCTGATCATACCTTTGTGTACTGATTTGCTCCATCGCCTTGTCTCGATTCAACCT 1628
Db |||||

Db 1101 TTTCAATTTAATCGTGCCATGTGTCTAATCTCATCGATGGCCCTACTGGGCTTCAATT 1160
Qy 1629 GCOCGAGATTGGGTGAAAAAATTATCGTGGGTGTTTACCATTCTTGTCTCTCGCTACCGT 1688
Db |||||
Qy 1161 GCCGCGGATTGGGCGAGAAACTGACGCTGGAGTTACAATTCTTCTATCGCTCACAGT 1220
Db |||||
Qy 1689 GTTCTGAATATGTTGCCGAGACAATGCCGGCTACTTCCGATGCGGTGGCAATTTGGAT 1748
Db |||||
Qy 1221 GTTCTCAACCTTTAGCTGAGACATTCGCCCAAGTATCTGATGCAATCCCTTTGTTAGG 1280
Db |||||
Qy 1749 ACGCATCGTGTGTTTGTGCTGCGCATGGCATATTCGAATGATGAGTGCCTCCAGGACGACC 1808
Db |||||
Qy 1281 CACTTACTTCAATTGATCATGTTTCAATGCTGCGCTGCGGTGGTGTGACAGTAGTGGT 1340
Db |||||
Qy 1809 GCTGATCCTAGATTCCCGACACCGCTGTTTCGAGACACATCTCTCCGAGCGGAAGCAC-- 1866
Db |||||
Qy 1341 GCTCAACTTACCACCATCGCACAGCGGACATTCACGAGATGCCACCGTGGATCAAGTCGT 1400
Db |||||
Qy 1867 ---CAGTACTCTCCGAGTTGAGCTGAAGAGCGCTGCTGCGAAATCGTCTGCGCCAAAC 1923
Db |||||
Qy 1401 TTTCTCAATGGCTGCCCTGGATCTTTCGGAATGGTGTGACCCCGTTCGCAAGATTACAG 1460
Db |||||
Qy 1924 GTACTAGACATCGATGATGACTTCCG-----GCACAATTTGTCGCCCATGAC 1970
Db |||||
Qy 1461 CAACAACATCTATTAAAGCATCCATGAAGAGCTGGAGCTAAAGAGCGCTCTCTCCAA 1520
Db |||||
Qy 1971 GCCGCGGGAACACTGCCACACACCGCGCTTTCTATCGCACGGTTTATGACAAAGCGGA 2030
Db |||||
Qy 1521 ATCCCTGCTGCCAATGTCTCGACATCGACGACGCTTCCGGCACACAATATCTGGGTC 1580
Db |||||
Qy 2031 CGATGGCAGCATTTGGGCAATTTGCGAGACCCGGAATGCGGATGCGGTGACCCATCATAC 2090
Db |||||
Qy 1581 CCAAAACCGCATTTGGCTCGTGGCCAGCTTGGTGGCCCAACAACGGTGGAGGAGCATCA 1640
Db |||||
Qy 2091 GTGCATCAATCATCAACTGAATATGAATAGTATTAACTTAAAGGAATTCGCTTTAT 2150
Db |||||
Qy 1641 CACGGCATCGGTGCAATCACAAGATCTTCACTCTAATTTCTCAAGAATTTGCAATTTAT 1700
Db |||||
Qy 2151 AACTGATCAGTACGTTAAAGATGACGAGTGCATGCAATTTGCCAATGATTGGAATTTGC 2210
Db |||||
Qy 1701 TAGCGCGGATGCGCAAGCTGACGACGAGGGAATTTGATCGCGGATTTGGAAGTTGCG 1760
Db |||||
Qy 2211 AGCTATGCTGTTGACAGACTGCGCTTATCATATTACAAATGTTCCGAATATTAGCAC 2270
Db |||||
Qy 1761 GGCATGTTGTGATAGATTTTGTAAATTTTTCACGCTCTTCAAGATTTATTGCAAC 1820
Db |||||
Qy 2271 AATAGCTGTACTACTATCGGCACCATATATTGT 2306
Db |||||
Qy 1821 GGTACGGTCTGCTCTCCGCTCGGCACATAATCGT 1856
Db |||||

RESULT 15
AF321447
LOCUS
DEFINITION
Drosophila melanogaster nicotinic acetylcholine receptor Dalphae
subunit variant type III (nAcRalpha-30D) mRNA, complete cds,
alternatively spliced.
AF321447
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1. (bases 1 to 2023)
Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalphae, Dalphae and Dalphae7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
Genetics 160 (4), 1519-1533 (2002)
21969411

PUBMED 11973307
REFERENCE 2 (bases 1 to 2023)
AUTHORS Grauso,M. and Sattelle,D.B.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
Genetics Unit, University of Oxford, South Park Road, Oxford OX1
3QX, UK

FEATURES
source Location/Qualifiers
1..2023
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="30D1"
/dev_stage="embryo"
1..2023
/gene="nAcRalpha-30D"
379..1863
CDS
/note="nAcRalpha-30D"
/gene="ion channel; neurotransmitter transmembrane
receptor; alternatively spliced; contains exons 3b and 8b"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalpha6 subunit
variant type III"
/protein_id="AA113394.1"
/db_xref="GI:20152849"
/translation="MDSLPASLSIFVLLIFLAIKESCQGPHEKRLNHLSTVNTL
ERPVANSEPEVRFGLTQQLIDVERKNQLITNMLSLWLNNDYLNLRNTERVGGVK
FLDITPQHKMFGSWTDGNQDLVNSDGLSDFTNGEWLLAMPKKNITVYA
CCPPYVDITETIQIRRTLYFFNLIVPCVLISMLLGLTLPDSEKLTGLVTIL
LSLTVNLVAETLPQVSDAIPLLGTIFENFMVASSWLTVVNLVYHRTADIEH
PMTKSVFLOMLPILMRGRPKTRKTILLNEMKELEKRSKSLANLVLDID
DPRHTISSQWLVAIGSSAFGRPTVEEHRTALGCKHKLHLILKELOPITARMKADD
BAELIGDWKFAAMVVDRELFVTLFTIIATVTVLLSAPHIIIVQ"
794
variation
/gene="nAcRalpha-30D"
/note="results in glycine to serine substitution; compared
to the sequence deposited in GenBank Accession
AE003626"
846
/replace="a"
variation
/gene="nAcRalpha-30D"
/note="results in asparagine to serine substitution;
compared to the sequence deposited in GenBank Accession
Number AE003626"
/replace="a"

ORIGIN
Query Match 19.1%; Score 441.6; DB 3; Length 2023;
Best Local Similarity 58.2%; Pred. No. 1.4e-109;
Matches 824; Conservative 0; Mismatches 574; Indels 18; Gaps 2;
QY 909 TAAAGTTTCCCTAGCAGATATCATGAAAGAGACTGTGTACAGATCTTTTGGATCCTTA 968
DB 441 TAAAGAAAGCTGTCAAGACCTCATGAAAGCGCTGTGTAACCATCTGCTGCCACCTA 500
QY 969 TAATACATAGAACCTCCGTTCTCAATGAATCGGACCGGTTCACATTAAGCTTTGGTTT 1028
DB 501 CAATACGCTGGAGCACCGCTGGCCCAATGAATCGGAGCCCTCGGAGGTTAAGTTCGGACT 560
QY 1029 AACTTTTAATGCAAAATTATCGATGTGACGAGAAAAATCAATTTGCTAGTCACTTAATGTGTG 1088
DB 561 GACGCTGCAGCAGATCATCGACGTGGACGAGAAGAAATCAGCTTCTCATTAACGAATCTTTG 620
QY 1089 GTTAAACATCGAGTGAACGACATGAATCTCCGTGGAAACACCTCCGACTATGGCGGAGT 1148
DB 621 GCTTTTCGTTGGAGTGAACGACTACAAATCTGGCTGGAATGAACGGAATACGGGGGGT 680
QY 1149 TAAGGATCTGCGAATACCCCGCATCGCATCTGGAAGCGGACGCTGCTGATGTACAAACAG 1208
DB 681 CAAGGATCTACGAATCACGCCCAACAGCTGTGGAGCCCGAGCTCATGTATCAACAG 740

Db 1821 GGTACGGTGTGCTTCCTCCGCGACATATCGT 1856

Search completed: May 8, 2004, 02:41:13
Job time : 9171.23 secs

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 10:08:00 ; Search time 929.055 Seconds
(without alignments)
10562.710 Million cell updates/sec

Title: US-09-303-232-1_COPY_372_2681

Perfect score: 2310

Sequence: 1 atgaaaaatgcacactgaa.....caccacatatattgtctcg 2310

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_29Jan04:*
1: Geneseqn1980s:*
2: Geneseqn1990s:*
3: Geneseqn2000s:*
4: Geneseqn2001as:*
5: Geneseqn2001bs:*
6: Geneseqn2002as:*
7: Geneseqn2003as:*
8: Geneseqn2003bs:*
9: Geneseqn2003cs:*
10: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	2310	100.0	2886	3	Aaz24475 D. melano
2	881	38.1	984	4	ABL27131
3	776.2	33.6	3144	4	ABL27130
4	531	23.0	803	4	ABL07799
5	512.8	22.2	3700	3	Aaz24476 H. viresc
6	479	20.7	936	4	ABL13733
7	412.2	17.8	1540	4	ABL07231
8	411.4	17.8	3109	3	Aaz24477
9	335.8	14.5	14688	4	ABL07798
10	296.2	12.8	1509	3	Aac58395
11	296.2	12.8	1509	4	AAC90380
12	296.2	12.8	1876	2	AAV12197
13	296.2	12.8	1876	2	AAT48239
14	296.2	12.8	1876	6	AB54875
15	296.2	12.8	1876	6	ABV73248
16	296.2	12.8	1876	8	ADAI0864
17	294.6	12.8	1509	4	AAC90385
18	294.6	12.8	1509	4	AAC90386
19	294.6	12.8	1590	2	AAV44687
20	293	12.7	1509	2	AAC90387
21	284.4	12.3	2101	2	AAT59197
22	283.8	12.3	1964	6	AB211298
23	267.8	11.6	2769	2	AAT59196

24	251	10.9	2277	2	AAV12199	Aav12199 Human neu
25	251	10.9	2277	2	AAQ90387	Aac90387 Alpha 2 s
26	251	10.9	2277	6	AB54870	Ab54870 Human neu
27	251	10.9	2277	6	ABV73243	Abv73243 Human neu
28	251	10.9	2664	2	AAT48235	Aat48235 Neuronal
29	251	10.9	2664	2	ABK92165	Abk92165 Prostate
30	251	10.9	2664	8	ADAI0854	Adai0854 Human neu
31	251	10.9	2664	9	ADC71170	Adc71170 Human 205
32	251	10.9	2666	9	ADC71168	Adc71168 Human 205
33	245.6	10.6	1416	4	AAC90382	Aac90382 Chimeric
34	242.6	10.5	1809	5	AAS91552	Aas91552 DNA encod
35	242.6	10.5	2082	2	AAT59527	Aat59527 Alpha4 su
36	242.6	10.5	2752	9	ADB78668	Adb78668 Human nic
37	242.6	10.5	2752	9	ADB78671	Adb78671 Human nic
38	242.6	10.5	2752	9	ADB78670	Adb78670 Human nic
39	242.6	10.5	2752	9	ADB78661	Adb78661 Human nic
40	242.6	10.5	2752	9	ADB78672	Adb78672 Human nic
41	242.6	10.5	2752	9	ADB78663	Adb78663 Human nic
42	242.6	10.5	2752	9	ADB78669	Adb78669 Human nic
43	242.6	10.5	2752	9	ADB78662	Adb78662 Human nic
44	242.6	10.5	2752	9	ADB78673	Adb78673 Human nic
45	242.6	10.5	3343	9	ADE85024	Ade85024 Farnesyl

ALIGNMENTS

RESULT 1
AAZ24475
ID AAZ24475 standard; CDNA to mRNA; 2886 BP.
AC AAZ24475;
XX
XX
DT 17-FEB-2000 (first entry)
XX
DE D. melanogaster acetyl-choline receptor DNA from clone Da7.
XX
KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
KW neurotransmission; plant protection agent; conductance; AChR; ds.
XX
OS Drosophila melanogaster.
XX
XX
PH Key Location/Qualifiers
CDS 372..2684
FT /*tag= a
FT /product= "acetyl choline receptor."
XX
PN DE19819829-Al.
XX
PD 11-NOV-1999.
XX
XX 04-MAY-1998; 98DE-01019829.
PR 04-MAY-1998; 98DE-01019829.
XX
PA (FARB) BAYER AG.
XX
PI Adamczewski M, Oellers N, Schulte T;
XX
XX WPI: 2000-014207/02.
DR P-PSDB; AAY50814.
XX
XX New nucleic acid encoding a nicotinic acetylcholine receptor from
PT insects, used to identify potential insecticides.
XX
XX Claim 1a; Page 8-12; 26pp; German.

This invention describes a novel nucleic acid (NA) encoding a nicotinic acetyl-choline receptor (I) from insects which can be used as an insecticide. Inhibitors of (I) interfere with neurotransmission (I) (also vectors containing it, its regulatory regions, and antibodies directed against (I)-encoded proteins) are used to screen for: (a) plant protection agents that alter conductance of AChR, potentially useful as

DR P-PSDB; ABB63696.

XX New isolated nucleic acid detection reagent for detecting 1000 or more

PT Genes from *Drosophila* and for elucidating cell signaling and cell-cell

PT interactions.

XX

XX Claim 1; SEQ ID NO 17879; 21pp + Sequence Listing; English.

XX

CC The invention relates to an isolated nucleic acid detection reagent

CC capable of detecting 1000 or more genes from *Drosophila*. The invention is

CC useful in developmental biology and in elucidating cell signalling and

CC cell-cell interactions in higher eukaryotes for the development of

CC insecticides, therapeutics and pharmaceutical drugs. The invention

CC discloses genomic DNA sequences (AB116176-AB130511), expressed DNA

CC sequences (AB101840-AB116175) and the encoded proteins (ABB57737-

CC ABB572072). The sequence data for this patent did not form part of the

CC printed specification, but was obtained in electronic format directly

CC from WIPO at ftp.wipo.int/pub/published_pct_sequences

XX

SQ Sequence 803 BP; 211 A; 199 C; 190 G; 203 T; 0 U; 0 Other;

Query Match 23.0%; Score 531; DB 4; Length 803;

Best Local Similarity 88.4%; Pred. No. 1.8e-138;

Matches 616; Conservative 0; Mismatches 15; Indels 66; Gaps 1;

QY 1054 GACGAGAAAATCAATTGCTAGTCACTAATGTGTGGTTAAACTGGAGTGAACGACATG 1113

Db 1 GACGAGAAAATCAATTGCTAGTCACTAATGTGTGGTTAAACTGGAGTGAACGACATG 60

QY 1114 AATCTCGCTGGAAACACCTCGCATATGCGGAGGTTAAGGATCTGCGAATACCCCGCAT 1173

Db 61 AATCTCGCTGGAAACACCTCGCATATGCGGAGGTTAAGGATCTGCGAATACCCCGCAT 120

QY 1174 CGCATCTGGAAGCGGACGCTCTGATGTACACAGTGGGATGAGGATTTGACGGCACC 1233

Db 121 CGCATCTGGAAGCGGACGCTCTGATGTACACAGTGGGATGAGGATTTGATGGCACC 180

QY 1234 TACCAGACGAACGCTGGTGGTGGCGAACAACGGCTCGTGTCTATACGTTCCGCGGGATC 1293

Db 181 TACCAGACGAACGCTGGTGGTGGCGAACAACGGCTCGTGTCTATACGTTCCGCGGGATC 240

QY 1294 TTCAAGTCGACGTCGAAGATCGACATCACGTGGTTCCTTCGATGACCGCGTGGAG 1353

Db 241 TTCAAGTCGACGTCGAAGATCGACATCACGTGGTTCCTTCGATGACCGCGTGGAG 300

QY 1354 ATGAAGTTCCGAGTTGGACCTACGACGATTCAGCTGGATTTACAATTACAAGATGA 1413

Db 301 ATGAAGTTCCGAGTTGGACCTACGACGATTCAGCTGGATTTCAAGTTTGGT----- 343

QY 1414 ACTGGCGGTGATATCAGCAGTTACGTGCTCAACGGCGAGTGGGAACCTACTGGTGGCC 1473

Db 344 -----TCAGTTGGCCC 354

QY 1474 GGCAAAAGTTAACGAGATCTATTCACTGTGCGCGGAACCTATATAGACATCACCTTC 1533

Db 355 GGCAAAAGTTAACGAGATCTATTCACTGTGCGCGGAACCTATATAGACATCACCTTC 414

QY 1534 GCCATCATATCCGCGAGCAACACTGTACTATTCTTCAACCTGATCATCTTGTTGA 1593

Db 415 GCCATCATATCCGCGAGCAACACTGTACTATTCTTCAACCTGATCATCTTGTTGA 474

QY 1594 CTGATTGCCCTTCCATGGCTTGTCTGGATTCACCTCGCGGACAGATTCGGTGAATAATTA 1653

Db 475 CTGATTGCCCTTCCATGGCTTGTCTGGATTCACCTCGCGGACAGATTCGGTGAATAATTA 534

QY 1654 TCCTGGGTGTTACCAATCTTGTCTCTCGTACCGCTGTTCTGAATATGTTGCCGAGACA 1713

Db 535 TCCTGGGTGTTACCAATCTTGTCTCTCGTACCGCTGTTCTGAATATGTTGCCGAGACA 594

QY 1714 ATGCCGGCTACTCCGATCGGTTGCCATTTGTGATAC 1750

Db 595 ATGCCGGCTACTCCGATCGGTTGCCATTTGTGATAC 631

RESULT 5

AAZ24476

ID AAZ24476 standard; cDNA to mRNA; 3700 BP.

XX

XX AC AAZ24476;

XX

DT 17-FEB-2000 (first entry)

XX

DE H. virescens acetyl-choline receptor DNA from clone Hva7-1.

XX

KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening; neurotransmission; plant protection agent; conductance; AChR; ds.

XX

OS Heliothis virescens.

XX

FH Key Location/Qualifiers

FT CDS 335..1825

FT /*tag= a

FT /product= "acetyl-choline receptor"

XX

DE19819829-A1.

XX

PD 11-NOV-1999.

XX

PF 04-MAY-1998; 98DE-01019829.

XX

PR 04-MAY-1998; 98DE-01019829.

XX

PA (PARB) BAYER AG.

XX

PI Adamczewski M, Oellers N, Schulte T;

XX

DR WPI; 2000-014207/02.

DR P-PSDB; AAY50815.

XX

PT New nucleic acid encoding a nicotinic acetylcholine receptor from insects, used to identify potential insecticides.

PT

XX Claim 1a; Page 14-17; 26pp; German.

XX

CC This invention describes a novel nucleic acid (NA) encoding a nicotinic acetyl-choline receptor (I) from insects which can be used as an insecticide. Inhibitors of (I) interfere with neurotransmission. (I) (also vectors containing it, its regulatory regions, and antibodies directed against (I)-encoded proteins) are used to screen for: (a) plant protection agents that alter conductance of AChR, potentially useful as insecticides, or (b) genes which encode polypeptides that are involved in formation of functionally related AChR in insects. (I) are also used to isolate and characterize the specified regulatory regions and for recombinant production of (II). This sequence encodes an acetyl-choline receptor isolated from *Heliothus virescens*

XX

SQ Sequence 3700 BP; 893 A; 953 C; 944 G; 910 T; 0 U; 0 Other;

Query Match 22.2%; Score 512.8; DB 3; Length 3700;

Best Local Similarity 62.5%; Pred. No. 5.5e-133;

Matches 878; Conservative 0; Mismatches 502; Indels 24; Gaps 4;

QY 925 GGATATCATGAAGAAGACGTGTACAGATCTTTGGATCTTTATATACACTAGAACGT 984

Db 425 GGGTACCAAGAGAGCGGCTACTGACACCTATTGGACCTACCAACCTGAGAGG 484

QY 985 CCCGTTCTCAATGAATCGGACCGGTTACAAATTAAGCTTTGGTTAACTTTATGCAATT 1044

Db 485 CCCGTCGTCAACGAGAGCGACCGCTGACGCTCTCTCCGGCTCAGCTCATGAGATC 544

QY 1045 ATCGATGTGGACGAGAAAATCAATGTCTAGTCACTAATGTGTGGTTAAAACTGGAGTGG 1104

Db 545 ATCGAGCTGGACGAGAGAACCAGCTTTTAAACAACATCTGGCTAAACTAGAGTGG 604

QY 1105 AACGATCAATCTCGGTGGAAACCTCCGACTATGCGGAGTTAAGGATCTGGAATA 1164

QY 1565 ATTCTTCAACCTGATCATACCTTGTGTAAGTGGCTTCCATGGCCTTCTCGGATTC 1624
|||
DB 766 ATTCTTCAATTAATGTGCCATGTGTCTAATCTCATGATGGCCCTTACTGGCTTCA 825
|||
QY 1625 CCCTGCCCGCAGATTCGGGTGAAAAATTATCGTGGGTGTTACCAATCTTCTCTCGTGA 1684
|||
DB 826 CATTCGCCCGGATTCGGCGAGAACTGACGCTGGCGTAATATATCTACTATCAATA 885
|||
QY 1685 CCGTGTCTTCAATATCGTTCGCCGAGACATGCGCGTACTTCGGATGCGGTGCC 1739
|||
DB 886 CAGTATTTCTAAACCTTGTGCCGAGTCCATGCCGACACGTCGGATGCTGTGCC 940
|||

RESULT 8

AAZ24477
ID AAZ24477 standard; cDNA to mRNA; 3109 BP.

XX AC

XX AC

XX AAZ24477;

DT 17-FEB-2000 (first entry)

XX DE

XX H. virescens acetyl-choline receptor DNA from clone Hva7-2.

XX DE

XX Acetyl-choline receptor; nicotinic; insect; insecticide; screening;

KW neurotransmission; plant protection agent; conductance; AChR; ds.

XX KW

XX Heliothis virescens.

XX OS

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

XX DE

QY 1044 TATCGATGTGGACGAGAAAAATCAATGCTAGTCACTAATGTGTGGTTAAAACTGGAGTG 1103
|||
DB 271 CATTCAGCTGGACGAGAGAAATCACTACTTATAACCAATATATGGCTGTCTGTTGGAGTG 330
|||
QY 1104 GAACGACATGAATCTCCGTGGAAACACCTCCGACTATGGCGGAGTTAAGATCTGCGAAT 1163
|||
DB 331 GAATGACTCAACCTGAGGTGGAACGACAGCGAGTATGGCGGGTCAAGGACCTTCAGGAT 390
|||
QY 1164 ACCGCCCATCGCATCTGGAGCGCGGACGCTGATGTACAACAGTCCGGATGAGGAT 1223
|||
DB 391 CAGCCCAACAAGTTGTGGAAGCCGGACGCTCTTATGTAATATAGTCTGACGAGGTTT 450
|||
QY 1224 TGACGGCACCTACACGACGAACTGTGTGGTGGGAAACAACGGCTCGTGTCTATACGTTC 1283
|||
DB 451 TGACGGGACCTACACGACCAACGCTGTGTGGTTCAGAAAGCGGCGGAGTTGCTGTACGTGCC 510
|||
QY 1284 GCCGGGATCTTCAAGTCGAGCTGCAAGATCGACATCAGCTGGTTCCTCCCTTCGATGACCA 1343
|||
DB 511 ACCTGCGATATTCAGAGACCATGCAAGATGGACATCGCGTGGTTTCCCTTCGACGACCA 570
|||
QY 1344 GCGTGGAGATGAAGTTCCGCACTTCGACGCGGATTCGAGCTGGATTTTCAAT 1403
|||
DB 571 ACATCTGATATGAAGTTCCGTAGCTGACATATGACGGCAATCAGTTGGATCTGGTGT 630
|||
QY 1404 ACAAGATGAACCTGGCGGTGATATCAGCAGTTAGTGTCAACGGCGAGTGGGAACTACT 1463
|||
DB 631 AAAAGATGAGCGAGCGGCGGATCTATCGGACTTCATAAATAATGGGAGTGGTATCTAAT 690
|||
QY 1464 GGGTGTCCCGGCAAACTAACGAGATCTATTACAACTGCTGCCCGGAAACCTATATAGA 1523
|||
DB 691 AGGAATGCCAGGCAAAAGAACACATAACATACCGTGTGCCCCGAGCCCTACGTGGA 750
|||
QY 1524 CATCACTTCGCCATCATCATCCCGCGACGAACTGTACTATTCTTCAACCTGATCAT 1583
|||
DB 751 CGTCACCTTCACCATCATGATGAAGAGAGAACCTTGTACTACTTCTTCAACCTGATCGT 810
|||
QY 1584 ACCTTGTGTACTGATTCCTCCATGGCTTCTCGGATTCACCTGCGCGCAGATTCGGG 1643
|||
DB 811 CCGTGGCTGTGATCTCATCGATGGCACTCTCGGCTTCACTGCGCCACGAGATCCGG 870
|||
QY 1644 TGAATAATATCGCTGGGTGTACCACTTCTGCTGCGTACCGTGGTCTGTAATATCGT 1703
|||
DB 871 AGAGAACTCACTTGGAGTCACTATTCTTATGCTGACGGTGTCTTCAACCTGCT 930
|||
QY 1704 TGCGGAGCAATGCCGCTACTTCCGATGCGGTGCCATGT 1744
|||
DB 931 AGCGAGACCTGCCACAGGTCTCCGAGCTATCCCCCTGT 971
|||

RESULT 9

ABL07798

ID ABL07798 standard; cDNA; 14668 BP.

XX AC

XX ABL07798;

XX AC

XX ABL07798;

DT 26-MAR-2002 (first entry)

XX DE

XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 17876.

XX DE

XX Drosophila; developmental biology; cell signalling; insecticide;

KW pharmaceutical; gene; ss.

XX KW

XX Drosophila melanogaster.

OS OS

XX OS

XX WO200171042-A2.

PN PN

XX PD

XX 27-SEP-2001.

XX PF

XX 23-MAR-2001; 2001WO-US009231.

XX PR

XX 23-MAR-2000; 2000US-0191637P.

PR 11-JUL-2000; 2000US-00614150.

QY 924 AGGATATCATGAAAGAGACTGTTACAGATCTTTTGGATCCTTATATACACTAGAAG 983
|||
DB 151 AGGTCTTCACGAGAGAGACTCTGTAACGCTGTGCTGGGAACTACACACCTCGAGCG 210
|||
QY 984 TCCGCTCTCAATGATCGACCGCTTCAATTAAGCTTTGTTAACTTTAATGCAAT 1043
|||
DB 211 ACCGGTGGCAACGAGAGGAAACCGCTAGAGTTCAGGTTCGGCTTGCAGCAAT 270
|||

Query Match

Best Local Similarity 17.8%; Score 411.4; DB 3; Length 3109;

Matches 565; Conservative 0; Mismatches 256; Indels 0; Gaps 0;

1091 TAAATCGAGTGGAGACACATGAATCTCCGCTGGAAACACCTCCGACTATGGCGAGTTA 1150
 233 TGCATATGCTTGGACAGATCACTATTTACAGTGAATGTGCAGATATATCCAGGGTGA 292
 1151 AGGATCTGCGAATACCCCGCATCGATCTGGAGCCGAGCGTGTCTGATGTACACAGTG 1210
 293 AGACTGTTCGTTTCCAGATGGCCAGATTTGGAAACACAGACATTTCTTCTATACAGTG 352
 1211 CGGATGAGGATTTGACGGACCTACAGACGAGACGTTGGTGGCGGAAACACCGCTCGT 1270
 353 CTGATGAGCGTTTGGACCCACATTCACACTAACCGTGTGGTGAATTTCTTCTGGGCA 412
 1271 GTCTATACGTTCCCGCGGATCTTCAAGTCGACGTCGAGATCGACATCACGTTGTTC 1330
 413 GCCATGACCTGCTCCAGGATATTCAGAGTCTCTGCTACATCGATGTACGTTGTTT 472
 1331 CCTTCGATACCGAGCGTGGAGATGAAGTTGGCAGTTGGACCTACGACGATTCGACG 1390
 473 CCTTTGATGTGACGACCTGCAAACTGAAGTTTGGTCTCTGCTTACGGAGGCTGTCT 532
 1391 TGGATTTACAAATACAGATGAAGTGGCGGTGATATCAGCAGTTACGTTCTCAACGCG 1450
 533 TGGATCTGCAGATCAGAGGCA-----GATATCAGTGGCTATATCCCAATGGAG 583
 1451 AGTGGGAATCTACTGGGTGTCCCGCAAAACGTAACGAGATCTATTAACAATGCTGCCCG 1510
 584 AATGGGACCTAGTGGGATCCCGGACGAGAGTGAAGGTTCTATGAGTGTGCAAG 643
 1511 AACCTATATAGACATCACCTTGGCATCATCATCGCGGACGACACTGTACTATTCT 1570
 644 AGCCCTACCCGATGTACCTTCAAGTACAGTACGCGCGGACGCTCTACTATGACC 703
 1571 TCACCTGATCATACCTTGTGTACTGATGCTGCTCCATGCGCTTGTGCGATTTCCTG 1630
 704 TCACCTGCTGATCCCTGTGTGCTCATCTCCGCGCTGCTGCTGCTGCTGCTGCTG 763
 1631 CGCCAGATTCGGGTGAAATATATCGTGGGTGTATCCATCTGCTCTGCTGACCGTGT 1690
 764 CTGACGATTCGGGAGAGATTTCCCTGGGATAACAGTCTTACTCTCTTACCGTCT 823
 1691 TTCTGAATATGGTTCGCGAGACATGCGCGTCTCTTCCGATGCGGTGCGCATTTG 1743
 824 TCATGCTGCTGCTGCTGAGATCATGCCGCAACATCCGATTCCGTTACCATTTG 876

RESULT 11
 AAC90380
 ID AAC90380 standard; cDNA; 1509 BP.

AC AAC90380;

XX 14-MAR-2001 (first entry)

DE Wild-type human alpha7 ligand gated ion channel coding sequence.

DE Human; alpha7 nicotinic acetylcholine gated ion channel;

KW 5-hydroxytryptamine; 5-HT3; calcium ion conductance; ss.

XX Homo sapiens.

OS WO200073431-A2.

XX 07-DEC-2000.

XX 25-MAY-2000; 2000WO-US011862.

XX 27-MAY-1999; 99US-0136174P.

XX (PHAA) PHARMACIA & UPJOHN CO.

XX Groppi VE, Wolfe ML, Berkenpas MB;

XX

DR WPI; 2001-061524/07.
 XX P-PSDB; AAB50012.
 PT Special cell culture medium for treating cells and for inducing mammalian
 PT cell lines to conduct calcium ions, comprising specified concentrations
 PT of ions of sodium, calcium and potassium at specified pH.
 XX Example 5; Page 60-61; 77pp; English.
 XX The present sequence is the coding sequence for wild-type human alpha7
 CC nicotinic acetylcholine gated ion channel. The human alpha7 ion channel
 CC was used in the construction of an alpha7/5-hydroxytryptamine (5-HT3)
 CC chimeric ligand gated ion channel (see AAC90382 and AAB50014). The
 CC alpha7/5-HT3 chimeric ion channel can be expressed by recombinant cells
 CC in the present invention, resulting in preferential calcium ion
 CC conductance by the cells
 XX Sequence 1509 BP; 297 A; 451 C; 429 G; 332 T; 0 U; 0 Other;

Query Match 12.8%; Score 296.2; DB 4; Length 1509;
 Best Local Similarity 60.7%; Pred. No. 2.9e-72;
 Matches 506; Conservative 0; Mismatches 318; Indels 9; Gaps 1;

Qy 911 AAGTTTCCTAGCAGGATATCATGAAAGAGACGTTTACAGATCTTTTGGATCCTATA 970
 Db 53 ACGTGTCTCTGCAAGCGAGTTCAGAGGAAGCTTTTACAAGAGCTGGTCAAGAACTACA 112
 Qy 971 ATACACTAGAACGTCCTGTTCTCAATGAATCGGACCGTTTACAATTAAGCTTTGGTTTAA 1030
 Db 113 ATCCCTTGGAGAGCGCGTGGCCAAATGACTCGCAACCACTCACCGTCTACTTCTCCCTGA 172
 Qy 1031 CTTTAATGCAAAATPATCGATGTGGACGAGAAATCAATTTGCTAGTCACTAATGTGGT 1090
 Db 173 GCCTCTTCAGATCATGACGTCGATGAGAAGAACCAAGTTTAAACCAACATTTGGC 232
 Qy 1091 TAAACTGGAGTGAAGACATGAATCTCCGCTGGAACACCTCCGACTATGCGCGAGTTA 1150
 Db 233 TGCAAAATGTCTTGACAGATCACTATTTACAGTGGAAATGTGTCAGATATCCAGGGTGA 292
 Qy 1151 AGGATCTGCGAATACCGCGCATCGCATCTGGAAGCGGACGCTGCTGATGTACAACTG 1210
 Db 293 AGACTGTGTTTCCAGATGGCCAGATTTGGAACACAGACATTTCTTCTATAACAGTG 352
 Qy 1211 CGGATGAGGATTTGACGGACCTTACAGACGAAAGTGTGTGTCGGGAAACAAACGCTCGT 1270
 Db 353 CTGATGAGCGCTTTGACGCCACATTCACACTAAACGTTTGGTGAATTTCTTCTGGCGCAT 412
 Qy 1271 GTCTATACGTTCCGCGGGATCTTCAAGTGCAGTGCAGATCGACATCGACATCGTGTTC 1330
 Db 413 GCCAGTACCTGCTCCAGGCATATTCAGAGTTTCTGCTCATCATGATGATGCTGCTGTTTC 472
 Qy 1331 CTTTCGATGACCGGTTGCGAGATGAAGTTTCGGCAGTTGGACCTTACGACGATTCACG 1390
 Db 473 CTTTGTGATGCGACATGCAAACTGAAGTTTGGTCTGCTGTTTACGAGGCTGCTCT 532
 Qy 1391 TGGATTTCAAAATCAAGATGAATACTGGCGGTGATATCAGCAGTTTACGTCACACGGG 1450
 Db 533 TGGATCTGAGATGACGAGGCA-----GATATCAGTGGCTATATCCCAATGAG 583
 Qy 1451 AGTGGGACTACTGGTGTCCCGCAAAACGTAACGAGATCTATTTACAACCTGCTGCCCG 1510
 Db 584 AATGGGACCTAGTGGGAATCCCGCAAGAGAGTGAAGGTTCTATGAGTGTGCAAG 643
 Qy 1511 AACCTATATAGACATCACCTTCCGCAATCATCOGCGGACGACACTGTACTATTCT 1570
 Db 644 AGCCCTACCCGATGTACCTTCAAGTACAGTACGCGCGGACGCTCTACTATGACC 703
 Qy 1571 TCACCTGATCATACCTTGTGTACTGATGCTGCTCCATGCGCTTGTGCGATTTCCTG 1630
 Db 704 TCACCTGCTGATCCCTGTGTGCTCATCTCCGCGCTGCTGCTGCTGCTGCTGCTG 763
 Qy 1631 CGCCAGATTCGGGTGAAATATATCGTGGGTGTATCCATCTGCTCTGCTGACCGTGT 1690
 Db 764 CTGACGATTCGGGAGAGATTTCCCTGGGATAACAGTCTTACTCTCTTACCGTCT 823
 Qy 1691 TTCTGAATATGGTTCGCGAGACATGCGCGTCTCTTCCGATGCGGTGCGCATTTG 1743
 Db 824 TCATGCTGCTGCTGCTGAGATCATGCCGCAACATCCGATTCCGTTACCATTTG 1690

Db 764 CTGCAGATTCGGGGAGAGATTTCCCTGGGGATAACAGCTCTTACTCTCTTACCCTCT 823
 QY 1691 TTCTGAATATGTTGGCCGAGACAAATGCGGCTACTTCCGATGCGGTGCCATTG 1743
 Db 824 TCATGCTGCTGTCGGCTGAGATCATGCCCGCAACATCCGATTCGGTACCAATTG 876

RESULT 12
 ID AAV12197 standard; cDNA; 1876 BP.
 AC AAV12197;
 XX
 XX 14-MAY-1998 (first entry)
 DT
 XX
 DE Human neuronal nicotinic acetylcholine receptor alpha-7 subunit cDNA.
 KW Human; neuronal nicotinic acetylcholine receptor; alpha-7 subunit;
 KW brain tissue; screening; NACHR; antibody; ds.
 XX
 XX Homo sapiens.
 OS
 PH Key Location/Qualifiers
 CDS 73..1581
 FT /*tag= a
 FT /product= "neuronal nicotinic acetylcholine receptor
 FT alpha-7 subunit"
 XX
 XX WO9420617-A2.
 XX
 XX 15-SEP-1994.
 XX
 XX 08-MAR-1994; 94WO-US002447.
 XX
 XX 08-MAR-1993; 93US-00028031.
 XX
 XX (SALK) SALK INST BIOTECHNOLOGY IND ASSOC.
 XX
 XX Elliott KJ, Ellis SB, Harpold MM;
 XX WPI; 1994-303024/37.
 DR P-PSDB; AAW44153.
 XX
 XX Human neuronal nicotinic acetylcholine receptor subunits and DNA - also
 XX transformed cells useful for screening cpds. which modulate activity of
 XX the receptor.
 PS Claim 8; Page 78-79; 99pp; English.
 XX
 CC The present sequence encodes a human neuronal nicotinic acetylcholine
 CC receptor (NACHR) subunit. The cells expressing the alpha and/or beta
 CC NACHR subunits may be used in a method of screening compounds to identify
 CC any which modulate the activity of human neuronal NACHR. Subunit specific
 CC antibodies may be used to monitor the distribution and expression density
 CC of various subunits in normal vs diseased brain tissues. Testing of
 CC single receptor subunits or specific receptor subunit combinations with a
 CC variety of potential agonists or antagonists provides information with
 CC respect to the function and activity of the individual subunits and
 CC should lead to the identification and design of compounds that are
 CC capable of very specific interaction with one or more receptor subtypes.
 CC The resulting drugs should exhibit fewer unwanted side effects than drugs
 CC identified e.g. screening with cells that express a variety of subtypes
 XX
 SQ Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;
 Query Match 12.8%; Score 296.2; DB 2; Length 1876;
 Best Local Similarity 60.7%; Pred. No. 3.3e-72;
 Matches 506; Conservative 0; Mismatches 318; Indels 9; Gaps 1;
 QY 911 AAGTTTGCCTAGCAGATATCATGAAGAAGACATGTTACAGATCTTTTGGATCCTTATA 970
 Db 125 ACCTGTCCTGCAAGCGGAGTTCAGAGGAAGCTTTACAGAGAGTGGTTCAGAACTACA 184

QY 971 ATACACTAGAACGTCCTCCGTTCTCAATGAATCGGACCCGTTTCAATTAAGCTTTGGTTTAA 1030
 Db 185 ATCCCTTTGGAGAGCCCTGGCCAAATGACTCGCAACCACTCACGCTCTACTTCTCCCTGA 244
 QY 1031 CTTTAAATGCCAAATTTATCGATGTGACGAGAAAATCAATTGCTAGTCACTAATGTGTGCT 1090
 Db 245 GCCTCTCTGCAGATCATGGACGTGGATGAGAAGAACCAAGTTTAAACCAACCAATTTGGC 304
 QY 1091 TAAAACTGGAGTGGACGACATGAATCTCCGCTGGAAACACCTCCGACTATATGGCGAGTTA 1150
 Db 305 TGCAAATGTCTTGGACAGATCACTATTATTACAGTGAATGTGTCAAGATATATCCAGGGTGA 364
 QY 1151 AGGATCTGCGAATACCCGCGCATCGATCTGGAAGCCGAGCTGCTGATGTACACAGTG 1210
 Db 365 AGACTGTTCCTTTCCAGATGGCCAGATTTGGAAACCAAGACATTTCTTCTCTATAACAGTG 424
 QY 1211 CGGATGAGGGAATTTGACGGCACCTTACCAGACGACGTGGTGGTGGCGGAACAACGCGTCTG 1270
 Db 425 CTGATGAGCGCTTTGACGCCACATTCACACTAACGTGTGGTGAATCTTCTGGGCATT 484
 QY 1271 GTCTATACGTTCCCGCGGGGATCTTCAAGTCGAGCTGCAAGATCGACATCACGTGGTTC 1330
 Db 485 GCCAGTACCTGCTCCAGGCATATTCAAGAGTTTCTCTCATATCGATGACGTGGTTC 544
 QY 1331 CCTTCGATGACCGCGTGGAGATGAAGTTTCGGCAGTTGGAACCTACGACGAGATTCGAGC 1390
 Db 545 CCTTGTATGTGCAGCACTGCAAACTGAAGTTTGGGTCTCTTACGGAGGCTGGTCT 604
 QY 1391 TGGATTTTACAATTACAAAGATGAAACTGGCGGTGATATACAGCAGTTACGTCTCAACGGCG 1450
 Db 605 TGGATCTGCAGATGCAGGAGGCA-----GATATCAGTGGCTATATCCCAATGGAG 655
 QY 1451 AGTGGGAATCTACTGGGTGTGCCCGCAAAAGTAAACGAGATCTATTACAATCTGTGCCCGG 1510
 Db 656 AATGGGACCTAGTGGGAATCCCGGCAAGAGGAGTGAAGGTTCTATGAGTGTGCTGCAAG 715
 QY 1511 AACCTATATAGACATCACCTTCGCCATCATATCCGCCCTCGCCCTGCTGGTGTCTGCTTCT 1570
 Db 716 AGCCCTACCCCGATGTACCTTCACAGTGACCATGCGCCGACGACGCTCTACTATGGCC 775
 QY 1571 TCAACCTGATCATACCTTTGTGTACTGATTCCTCCATGGCTTGTCTCGGATTCACCCCTGC 1630
 Db 776 TCAACCTGTGATCCCTGTGTGCTCATCTCCGCCCTCGCCCTGCTGGTGTCTGCTGCTTC 835
 QY 1631 CGCCAGATTCGGTGAATAAATTATCGCTGGGTGTGTACCATCTTGTCTCTCGCTGACCCGTG 1690
 Db 836 CTGCAGATTCGGGGAGAAATTTCCCTGGGGATAACAGTCTTACTCTCTCTTACCGTCT 895
 QY 1691 TTCTGAATGTTGGCGAGACAATGCGGGCTACTTCCGATGGGTGGCCATTG 1743
 Db 896 TCATGCTGCTCGTGGCTGAGATCATGCCGCAACATCCGATTCGTACCAATTG 948

RESULT 13

AA48239
 ID AA48239 standard; DNA; 1876 BP.
 XX
 XX AAT48239;
 AC
 XX
 XX
 DT 09-APR-1997 (first entry)
 XX
 DE Neuronal nicotinic acetylcholine receptor alpha-7 subunit DNA.
 XX
 KW Neuronal nicotinic acetylcholine receptor; nAChR; neurotransmitter;
 KW ligand-gated receptor; ds.
 XX
 OS Homo sapiens.
 XX
 PH Key Location/Qualifiers
 CDS 73..1581
 FT /*tag= a
 XX
 XX WO9641876-A1.

XX PD 27-DEC-1996.
XX PF 07-JUN-1996; 96WO-US009775.
XX PR 07-JUN-1995; 95US-00484722.
XX PA (SIBI-) SIBIA NEUROSCIENCES INC.
XX PI Elliott KJ, Harpold MW;
XX WPI; 1997-065463/06.
XX P-PSDB; AAW09025.
XX Nucleic acids encoding nicotinic acetylcholine receptor sub-units - used
PT in screening to determine the effect of drugs on the receptor.
XX PS
XX PS Disclosure; Page 71-73; 108pp; English.
XX CC A DNA sequence (AAU48239) codes for the alpha-7 subunit (AAW09025) of the
CC human neuronal nicotinic acetylcholine receptor (nAChR). Host cells, esp.
CC mammalian cells or amphibian oocytes, carrying alpha-7 nucleic acids,
CC opt. in combination with other alpha and/or beta subunit nucleic acids
CC (see also AAU48232-38, AAU48240-41), express recombinant nAChR subunits
CC useful for identifying cpds. that modulate the activity of human nAChRs
XX CC
XX SQ Sequence 1876 BP; 369 A; 553 C; 531 G; 423 T; 0 U; 0 Other;
Query Match 12.8%; Score 296.2; DB 2; Length 1876;
Best Local Similarity 60.7%; Pred. No. 3.3e-72;
Matches 506; Conservative 0; Mismatches 318; Indels 9; Gaps 1;
QY 911 AAGTTTCCCTAGCAGGATATCATGAAGAAGAGACTGTTACAGATCTTTTGGATCCTTATA 970
DB 125 ACGTGTCTCCCTCAAGGCGAGTTCAGAGAGAGCTTTCAAGAGAGCTGTCAAGAACTACA 184
QY 971 ATACACTAGAACGTCCTTCTCAATGAATCGGACCCGCTTACAAATTAAGCTTTGGTTAA 1030
DB 185 ATCCCTTGGAGAGGCCGTGGCCCAATGACTCGCAACCACTCACCGTCTACTTCTCCCTGA 244
QY 1031 CTTTAAATGCAATATATGATGACGAGAGAAATCAATTCCTAGTCTACTAATGTGTTGGT 1090
DB 245 GCCTCTCTGCAATGATGACGCTGGATGAGAGAACCAAGTTTAAACCAACCAATTTGGC 304
QY 1091 TAAACTGGAGTGAACACATGATCTCCGCTGGAACACCTCCGACTATGCGGAGTTA 1150
DB 305 TGCATAATGCTTGGACAGATCACTATTTACAGTGAATGTGTCAAGATATCCAGGGTGA 364
QY 1151 AGGATCTCGGAATACCGCCGATCGCATCTGGAAGCCGCGAGTGTGTATGATACACAGTG 1210
DB 365 AGACTGTTCGTTCCAGATGGCCAGATTTGGAAACACAGACATTTCTTCTATAACAGTG 424
QY 1211 CGGATGAGGATTTGACGGACCTTACAGAGAACGTTGGTGGTGGGAAACACGCTCGT 1270
DB 425 CTGATGAGCGCTTTGACGCCACATTTCCACACTAACGTTGTGTGAATTTCTTCTGGGCATT 484
QY 1271 CTCATATAGTTCGCGCGGATCTTCAAGTGCAGTCAAGATCGACATCGAGTGTTC 1330
DB 485 GCCAGTCTCGCTCCAGGCAATTTCAAGATTTCTGCTACATCATGATGATGCTGTTTC 544
QY 1331 CTTTCGATGACAGCGGTGCGAGATGAAGTTTCGGCAGTTGGACCTTACGACGATTTCCAGC 1390
DB 545 CTTTGTATGTGACGACTGCAAACTGAAGTTTGGGTCTGTTTACGGAGGCTGGTCT 604
QY 1391 TGGATTACAATTAAGATGAACTGCGGTGATATCAGCATTTACGTGTCAACGGG 1450
DB 605 TGGATCTGAGATGACGAGGCA-----GATATCAGTGGCTATATCCCAATGAG 655
QY 1451 AGTGGAACTACTGGGTGTGCCCGCAACACGTAACAGATCTTATTAACACTGCTCCCGG 1510
DB 656 AATGGACCTAGTGGGATCCCGGCAAGAGAGTGAAGGTTCTATGATGCTGCAAG 715
QY 1511 AACCCCTATATAGACATCACCTTTCGCCATCATCATCCGCCGAGAACACTGTACTATTCTT 1570

DB 716 AGCCCTACCCCGATGTCACTTCCAGTGACCATGCCGCGCAGACCTCTACTATGCC 775
QY 1571 TCAACCTGATCATACCTTGTGTACTGATTCCTCCATGGCTTGTCTCGGATTCACCCCTGC 1630
DB 776 TCAACCTGCTGATCCCTCTGTGTCTATCTCCGCCCTCGCCCTGCTGTCTCTGCTTC 835
QY 1631 CGCCAGATTCGGGTGAAAAATATCGCTGGGTGTACCAATCTTGTCTCGCTGACCGTGT 1690
DB 836 CTGCAGATTCGGGGAGAGATTTCCTCGGGGATAACAGTCTTACTCTCTTACCGTCT 895
QY 1691 TTCTGAATATGTTGCCGAGACAAATGCCGCTACTTCCGATCGGTTGCCATTG 1743
DB 896 TCATGCTGCTGCTGCTGATCATGATCGCCGCAACATCCGATTCGTTACCATTTG 948
RESULT 14
ABS54875
ID ABS54875 standard; cDNA; 1876 BP.
XX AC ABS54875;
XX AC ABS54875;
XX DT 06-DEC-2002 (first entry)
XX DE Human neuronal nicotinic acetylcholine receptor alpha 7 subunit cDNA.
XX KW Human; neuronal nicotinic acetylcholine receptor; nNACHR; gene; ss;
XX KW ion flux; alpha 7 subunit.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 73..1581
XX FT /*tag= a
XX FT /product= "Human nNACHR alpha 7 subunit"
XX FN US6440681-B1.
XX PD 27-AUG-2002.
XX PF 07-JUN-1995; 95US-00487596.
XX PR 03-APR-1990; 90US-00504455.
XX PR 30-NOV-1992; 92US-00938154.
XX PR 08-MAR-1993; 93US-00028031.
XX PR 08-NOV-1993; 93US-00149503.
XX PA (MERI) MERCK & CO INC.
XX PI Elliott KJ, Ellis SB, Harpold MW;
XX WPI; 2002-711528/77.
XX P-PSDB; ABS70492.
XX DR
XX DR Identifying antagonists or agonists of human neuronal nicotinic
PT acetylcholine receptors, by contacting recombinant cells with test
PT compound, and measuring ion flux of cells or binding of compound to
XX nNACHR.
XX Claim 101; Col 57-60; 56pp; English.
XX The invention relates to a method for identifying compounds that are
CC antagonists or agonists of human neuronal nicotinic acetylcholine
CC receptors (nNACHRs), by contacting recombinant cells with a test compound
CC and measuring ion flux, the electrophysiological response of the cells or
CC binding of the test compound to the nNACHR. The recombinant cells are
CC produced by transfection with a nucleic acid encoding at least one human
CC nNACHR (alpha or beta) subunit, such that the cells express an nNACHR
CC comprising one human subunit encoded by the transfected nucleic acid.
CC This sequence represents cDNA encoding the alpha 7 subunit of the human
XX nNACHR polypeptide
XX SQ Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;

Query Match 12.8%; Score 296.2; DB 6; Length 1876;
Best Local Similarity 60.7%; Pred. No. 3.3e-72;
Matches 506; Conservative 0; Mismatches 318; Indels 9; Gaps 1;

QY 911 AAGTTTGCTAGCAGGATATCATGAAAGAGAGACTGTTACACGATCTTTGGATCCTTATA 970
Db 125 ACGTGTCCCTGCAAGCGAGTTCAGAGGAGCTTTTACAGGAGCTGGTCAAGACTACA 184
QY 971 ATACACTAGAACGTCCTCCCTCTCATGATCGGACCCGTTCAATTAAGCTTTGGTTTAA 1030
Db 185 ATCCCTTGGAGAGCCCTGGCCAAATGACTCGCAACCACTCACCGTCTACTTCTCCCTGA 244
QY 1031 CTTTAATGCAATATTCGATGTGGACGAGAAAAATCAATTGCTACTATGTGTGT 1090
Db 245 GCCTCTCGAGATCATGACGTGGATGAGAGAACCAAGTTTAAACCAACATTTGGC 304
QY 1091 TAAAACTGGAGTGGAAACGACATGATCTCCCTCGAACACCTCCGACTATGCGGAGTTA 1150
Db 305 TSCAAATGCTTTGGACAGATCACTATTACAGTGGATGTGTCAGAAATATCCAGGGTGA 364
QY 1151 AGGATCTGGGAATACCGCGCATCGCATCTGGAAGCCGAGCTGTGTGATGATACACAGTG 1210
Db 365 AGACTGTTCGTTCCAGATGCCAGATTTGGAAACCAAGACATTTCTCTATAACAGTG 424
QY 1211 CGGATGAGGGATTGACGGCACTTACCAGACGAGCGTGGTGGCGGAACAACGCTCGT 1270
Db 425 CTGATGAGCGTTTGACGCCACATCCACATTAAGTGTGGTGAATTTCTCGGGCAT 484
QY 1271 GTCTATAGTTCGCCCGGGATCTTCAAGTCGAGTCGACATCGATCGATCGTGTTC 1330
Db 485 GCCAGTACCTGCTCCAGGCATATCAAGATTTCTGCTACATGATGATCGTGTTC 544
QY 1331 CTTTCGATGACCGGTCGAGATGAGTTTCGCGAGTTGACCTAGACGGATTCGAGC 1390
Db 545 CCITTGATGTCAGCATGTCGAACTGAAATTTGGTTCCTGCTTACGGAGCTGGTCT 604
QY 1391 TGGATTTTACAATTAACAAGATGAACCTGCGGTGATATCAGCAGTTACGTGCTCAACGCG 1450
Db 605 TGGATCTGCAGATCGCAGGCA-----GATATCAGTGGCTATATCCCAATGAG 655
QY 1451 AGTGGGAATCTAGTGGTGTGCGCGGAAAGTAAACGAGATCTATTAACATGCTGCGCGG 1510
Db 656 AATGGGACCTAGTGGGAATCCCGGCAAGAGGAGTGAAGGTTCTATGAGTGTGCAAG 715
QY 1511 AACCTATATAGACATCACCTTCGCATCATCTCCGCGAGCAACACTGATCTATTCT 1570
Db 716 AGCCCTACCCCGATGTCACCTTACAGTGAACATGCGCCGAGACGCTCTACTATGGC 775
QY 1571 TCAACCTGATCATACCTTGTGTACTGATTCCTCCATGGCTTGTCTCGGATTCACCTGC 1630
Db 776 TCAACCTGCTGATCCCTGTGTCTCATCTCCGCTTGTGCTGCTGCTGCTGCTGCTGCT 835
QY 1631 CGCAGATTCGGTGAAAAATATTCGCTGGGTGTTACCATCTTCTGCTCGCTGACCGTGT 1690
Db 836 CTGCAATTCGGGGAGAAATTTCCCTGGGGATAACAGTCTTACTCTCTTTACCGTCT 895
QY 1691 TTCTGAATATGTTGCGAGACATCCGCTACTTCCGATGGGTGCGCATTTG 1743
Db 896 TCATGCTGCTGTTGATGATGCGCCCAACATCCGATTCGATCCGATTCGATCCGATTCG 948

RESULT 15
ABV73248

ID ABV73248 standard; cDNA; 1876 BP.

XX AC ABV73248;

XX DT 22-JAN-2003 (first entry)

XX DE Human neuronal NACHR alpha7 subunit encoding cDNA.

XX KW Human; neuronal; nicotinic acetylcholine receptor; NACHR; drug screening;

immunochemistry; NACHR alpha7 subunit; gene; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
XX CDS 73..1581
FT /*tag= a
FT /product= "NACHR alpha7 subunit"
FT /note= "neuronal nicotinic acetylcholine receptor"
XX WO200259266-A2.
XX PD 01-AUG-2002.
XX PF 29-OCT-2001; 2001WO-US050985.
XX PR 01-NOV-2000; 2000US-00703951.
XX PA (MERI) MERCK & CO INC.
XX PI Gillespie A, Claeps BO, Chavez-Noriega LE, Siegel R, Elliott KJ;
XX WPI; 2002-698532/75.
XX DR P-PSDB; ABB82435.
XX CC Cell comprising nucleic acids encoding human alpha and beta subunits of
XX PT neuronal nicotinic acetylcholine receptors, useful for in vitro screening
XX PT of a drug substance in a test system specific for humans.
XX PS Example; Page 129-130; 143pp; English.
XX CC The invention relates to a suitable host cell transfected with an
XX CC isolated nucleic acid molecule comprising a sequence of nucleotides or
XX CC ribonucleotides that encodes at least one alpha or beta subunit of a
XX CC human neuronal nicotinic acetylcholine receptor (NACHR). The compositions
XX CC and methods of the present invention, which provide a means to prepare
XX CC synthetic or recombinant receptors and receptor subunits that are
XX CC substantially free of contamination from many other receptor proteins,
XX CC are useful for observing the effect of a drug substance on a particular
XX CC subtype to perform in vitro screening of the drug substance in a test
XX CC system that is specific for humans. The antibodies can be used in
XX CC immunochemistry and for diagnostic and therapeutic applications. The
XX CC present sequence represents a human neuronal NACHR alpha7 subunit
XX CC encoding cDNA

SQ Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;

Query Match 12.8%; Score 296.2; DB 6; Length 1876;
Best Local Similarity 60.7%; Pred. No. 3.3e-72;
Matches 506; Conservative 0; Mismatches 318; Indels 9; Gaps 1;

QY 911 AAGTTTGCTAGCAGGATATCATGAAAGAGACTGTTACACGATCTTTGGATCCTTATA 970
Db 125 ACGTGTCCCTGCAAGCGAGTTCAGAGGAGCTTTTACAGGAGCTGGTCAAGACTACA 184
QY 971 ATACACTAGAACGTCCTCCCTCTCATGATCGGACCCGTTCAATTAAGCTTTGGTTTAA 1030
Db 185 ATCCCTTGGAGAGCCCTGGCCAAATGACTCGCAACCACTCACCGTCTACTTCTCCCTGA 244
QY 1031 CTTTAATGCAATATTCGATGTGGACGAGAAAAATCAATTGCTACTATGTGTGT 1090
Db 245 GCCTCTCGAGATCATGAGCTGGAGCTGGATGAGAAACCAAGTTTAAACCAACATTTGGC 304
QY 1091 TAAAACTGGAGTGGAAACGACATGATCTCCGCTGGAAACCTCCGACTATGGCGAGTTA 1150
Db 305 TSCAAATGCTTTGGACAGATCACTATTACAGTGGATGTGTCAGAAATATCCAGGGTGA 364
QY 1151 AGGATCTGGAATACCGCGCATTCGCACTGGAAGCCGAGCTGTGTGATGATACACAGTG 1210
Db 365 AGACTGTTCGTTCCAGATGGCCAGATTTGGAAACCAAGACATTTCTCTATAACAGTG 424
QY 1211 CGGATGAGGGATTGACGGCACTTACCAGACGATCGATCGGAGCAACGCTGTGCGGATTCGCT 1270

Db 425 CTGATGAGCGCTTTGAGCCACATTCACACTAACGTTGGTGAATTTCTTGGGCATT 484
Qy 1271 GTCTATACGTTCCGCCGGGATCTTCAAGTCGAGTCGACATCAGTGGTTCC 1330
Db 485 GCCAGTACCTGCTCCAGGCATATTCAAGAGTTCTGTCTACATCGATGTAGCTGGTTTC 544
Qy 1331 CCTTCGATGACCGGCTGCGAGATGAAGTTGCGCAGTTGGACCTACGACGGATTCCAGC 1390
Db 545 CCITTGATGCGACACTGCAAACTGAAGTTTGGTCTGTGTCCTACGGAGCTGGTCCT 604
Qy 1391 TGGATTTACAAATTACAAAGATGAAGTGGCGGTGATATCAGCAGTTACGTGCTCAACGGCG 1450
Db 605 TGGATCTGCAGATCAGGAGCA-----GATATCAGTGGCTATATCCCAATGGAG 655
Qy 1451 AGTGGGAACCTACTGGGTGTGCCCGGCAACGTAACGAGATCTATTACACTGCTGCCCGG 1510
Db 656 AATGGGACCTAGTGGGAATCCCGGCAAGAGGAGTGAAGGTTCTATGAGTGTGCAAAAG 715
Qy 1511 AACCTTATAGACATCACCTTCGCCATCATCCGCCGACGAACTGTACTATTCT 1570
Db 716 AGCCCTACCCGATGTACCTTCACAGTGACCATGGCCGCGAGGACGCTCTACTATGGCC 775
Qy 1571 TCAACCTGATCATACCTTTGTACTGTGCTCCATGCGCCTTGCTCGGATTCACCTGC 1630
Db 776 TCAACCTGCTGATCCCTGTGTCTCATCTCCGCCCTCGCCCTGTGGTGTCTCTGCTTC 835
Qy 1631 CGCCAGATTCGGGTGAATAATTATCGCTGGGTGTACCATCTTGTCTCGCTGACCGTGT 1690
Db 836 CTGCAGATTCGGGGAGAAGATTTCCCTGGGGATAACAGTCTTACTCTCTTACCGTCT 895
Qy 1691 TTCTGAATATGTTGCCGAGACAATGCCGGCTACTTCCGATGCGGTGCCATTG 1743
Db 896 TCATGCTGCTGCTGAGATCATGCCGCAACATCCGATTCGATCCATTG 948

Search completed: May 7, 2004, 15:01:30
Job time : 939.055 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 14:08:32 ; Search time 6207.93 Seconds
(without alignments)

11111.850 Million cell updates/sec

Title: US-09-303-232-1_COPY_372_2681

Perfect score: 2310

Sequence: 1 atgaagaatgcacaaactgaa.....caccacatatattgtctcg 2310

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_est1:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
C 1	424.8	18.4	885	12	BG632919	BG632919 GH16126.3
2	332.2	14.4	1101	29	CNS000IF	AL058211 Drosophil
3	331.2	14.3	1007	29	CNS00HJU	AL073676 Drosophil
4	323.8	14.0	978	29	CNS006F9	AL064281 Drosophil

5	281.8	12.2	1201	9	AL530299	AL530299
6	265	11.5	607	9	AI292581	AI292581 GH15518.5
7	251	10.9	2296	14	CD013901	CD013901 90134548
8	235.2	10.2	922	13	BU915857	BU915857 AGENCOURT
9	232.6	10.1	2940	11	AK034228	AK034228 Mus muscu
10	232.6	10.1	3230	11	AK083157	AK083157 Mus muscu
11	228.8	9.9	833	14	CB245337	CB245337 UI-M-FY0-
12	225.4	9.8	1436	29	AY402873	AY402873 Homo sapi
13	222.2	9.6	1436	29	AY402874	AY402874 Pan trogl
14	219.6	9.5	755	13	BU702422	BU702422 UI-M-FI0-
15	219.2	9.5	1374	29	AY406232	AY406232 Mus muscu
16	219.2	9.5	4290	11	AK029177	AK029177 Mus muscu
17	218.6	9.5	730	14	CF742344	CF742344 UI-M-HB0-
18	217	9.4	615	14	CB149460	CB149460 K-RST0205
19	215.8	9.3	1436	29	AY402875	AY402875 Mus muscu
20	215.8	9.3	1864	11	AK053497	AK053497 Mus muscu
21	215.8	9.3	2916	11	AK051730	AK051730 Mus muscu
22	215.8	9.3	3126	11	AK080415	AK080415 Mus muscu
23	207	9.0	3483	11	AK081254	AK081254 Mus muscu
24	207	9.0	4037	11	AK049722	AK049722 Mus muscu
25	207	9.0	4046	11	AK051742	AK051742 Mus muscu
26	206.4	8.9	683	12	BM639954	BM639954 170006876
27	205.4	8.9	908	13	BU149265	BU149265 AGENCOURT
28	203.4	8.8	902	29	AY407186	AY407186 Mus muscu
29	202.8	8.8	411	12	BI516733	BI516733 B160023A
30	202.2	8.8	2010	12	BI516843	BI516843 B160023B
31	201.8	8.7	410	12	BM711715	BM711715 UI-E-CL1-
32	200	8.7	658	12	BM711715	BM711715 UI-E-CL1-
33	199.6	8.6	902	29	AY407184	AY407184 Homo sapi
34	198	8.6	755	10	AW914206	AW914206 EST345510
35	197.2	8.5	1036	13	EX437801	EX437801 BX437801
36	196.8	8.5	1374	29	AY406230	AY406230 Homo sapi
37	196.8	8.5	1466	14	CD013891	CD013891 90139069
38	196.6	8.5	2513	11	AK033068	AK033068 Mus muscu
39	194.4	8.4	1603	29	AY411327	AY411327 Homo sapi
40	194	8.4	759	14	CD804155	CD804155 UI-M-GV0-
41	191.4	8.3	791	14	CF739825	CF739825 UI-M-HD0-
42	189.6	8.2	720	14	CA373069	CA373069 647093 NC
43	188.2	8.1	1374	29	AY406231	AY406231 Pan trogl
44	186.4	8.1	677	14	CF919760	CF919760 Bf1or531.
45	183.4	7.9	1442	29	AY402876	AY402876 Homo sapi

ALIGNMENTS

RESULT 1
BG632919/c
LOCUS
DEFINITION
GH16126.3prime GH Drosophila melanogaster head pOT2 Drosophila
melanogaster cDNA clone GH16126 3 similar to CG4128; Fban0004128
, ion channel', located on: 2L 30D1-30E1;: 04/10/2001, mRNA sequence.
ACCESSION
BG632919
VERSION
BG632919.1 GI:13758409
KEYWORDS
EST.
SOURCE
Drosophila melanogaster (fruit fly)
ORGANISM
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera, Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE
1 (bases 1 to 885)
Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G.,
Lewis,S. and Rubin,G.M.
BDGP/HMI Drosophila EST Project
Unpublished (2001)
Other ESTs: GH16126.5prime
Contact: Stapleton, M.
BDGP
Lawrence Berkeley National Lab
One Cyclotron Rd, Berkeley, CA 94720, USA
Fax: 510 486 6798
Email: http://www.fruitfly.org/EST_estafruitfly.berkeley.edu
Based upon the presence of a XhoI site followed by a run of 14 or

more T residues at the beginning of the sequence, this clone was polyadenylated. The resulting Poly-T sequence has been removed. hit genomic AE003511: arm:x [18792641,19136447] estimated-cyto:1BA3-18C6: 04/10/2001 Plate: GH.161 row: C column: 2 High quality sequence stop: 784.

FEATURES

Location/Qualifiers

source

1. .885

/organism="Drosophila melanogaster"

/mol_type="mRNA"

/db_xref="taxon:7227"

/clone="GH16126"

/sex="male and female"

/dev_stage="adult"

/lab_host="DH5 - alpha"

/clone_lib="GH Drosophila melanogaster head p072"

/note="Organ: head; Vector: p072; Site: 1; EcoRI; Site: 2;

XhoI; Sized fractionated cDNAs were directly ligated into

p072. Plasmid cDNA library."

ORIGIN

Query Match 18.4%; Score 424.8; DB 12; Length 885;
Best Local Similarity 74.6%; Pred. No. 1.8e-92;
Matches 534; Conservative 0; Mismatches 182; Indels 0; Gaps 0;

QY 1028 TAACCTTAATGCAATATTCGATGTGGACGAAATCAATTCGTAGTCACTAATGCT 1087
DB 884 TAACACTCATGCAATATTCGATGTGGACGAAATCAATTCGTAGTCACTAATGCT 825

QY 1088 GGTAAACCTGAGTGGAAACACATGAATCTCCGCTGGAAACACCTCCGACTATGCGGAG 1147
DB 824 GGCTCAATGGAATGGAACCATATGAATCTTCGATGGAATTCGATGAGTTCGGTGGTG 765

QY 1148 TTAAGGATCTCGAATACCGCGCATCGCATCTGGAACCGGACGCTGTGATGTACAACA 1207
DB 764 TCGCGGATCTCGAATTCGCGCACATCGCTATGGAACCGGATCTACTGTATGACAACA 705

QY 1208 GTGCGGATGAGGATTTGACGGCACTACAGACGACGCTGTGTGCGGAACACGCT 1267
DB 704 GTGCGGACGAGGGCTTCGATGGAACGAGTACGCCAANAATGTGTGTTCGCAATATGGGA 645

QY 1268 CGTGTCTATAGCTTCGCGCGGGGATCTTCAAGTCGACGTCGAAGATCGACATCAGTGGT 1327
DB 644 GCTGTCTAGCTACGCGCCAGGTATTTAAGTCAACGTTGAAGATCGACATTCAGTGGT 585

QY 1328 TCCCTTCGATGACAGCGGTGCGAGATGAAGTTCGGCAGTGTGGACCTACGACGATTC 1387
DB 584 TTCCATTCGACGATCAGAGATGTGAAATGAAATTTGGTTCGTGGACCTACGATGGGTTTC 525

QY 1388 AGCTGGATTTACAATTACAAGATGAATGCGCGGTGATATCAGCAGTTACGTCTCAACG 1447
DB 524 AGTTGACCTGAGTTGAGGACGAGAGTGTGGCGGACATTTCTAGCTTTATACCAATG 465

QY 1448 GCGAGTGGGAUACTACTGGGTGTGCCCGCAACGTPAACGAGATCTTATTAACAATCTGTC 1507
DB 464 GCGAATGGGACTTGTAGTGTGCCCGGTAAACGAAATGAAATCTACTATAATTTGCTGCC 405

QY 1508 CGGAACCTTATAGACATCACCCTTCGCCATCATCATCGGCCGACGACACTGTACTATT 1567
DB 404 CAGAACCTTTATATTGACATAACATTTGCGCAATTTGATTAAGGCGCAAAAGTTGTACTATT 345

QY 1568 TCTTCAACCTGATCATACCTTGTACTGATTTGCTTCCATGCGCTTGTCTCGGATTTCAACC 1627
DB 344 TTTTCAATCTGATGTGCGGTGCGTACTGATCGCTTCCATGCGACTGCTAGGTTTACAC 285

QY 1628 TCCCGCCAGATTCGGGTGAAAAATTATCGTGGGTGTTACCAATCTTGTCTCTCGCTGACCG 1687
DB 284 TCCGACCAAGATTCGTGTGAAAAGCTTTCGTTGGAGTTACAATTTCTATTATCGCTTACAG 225

QY 1688 TGTTCCTGAATGTTGCGGAGACATGCGCGGCTACTTCCGATCGGTGCGCATTTG 1743
DB 224 TCTTCTCAACATGTTGGCGGAAACATGCGCGGACCTCCGATGCGGTACCGCTG 169

RESULT 2

CNS00001F

LOCUS

DEFINITION

CNS00001F 1101 bp DNA linear GSS 03-JUN-1999
Drosophila melanogaster genome survey sequence T7 end of BAC #
BACR02C08 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

TITLES

AUTHORS

JOURNAL

COMMENT

Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : seqref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)

Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see <http://www.fruitfly.org> The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osogawa and
Aaron Mammoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain Y2; cn bw sp, the same strain used for the BDGP's
p1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.

FEATURES

source

1. .1101

/organism="Drosophila melanogaster"

/mol_type="genomic DNA"

/db_xref="taxon:7227"

/clone="BACR02C08"

/clone_lib="RPCI-98"

/note="end : T7"

ORIGIN

Query Match 14.4%; Score 332.2; DB 29; Length 1101;
Best Local Similarity 98.5%; Pred. No. 8.2e-70; Indels 0; Gaps 0;
Matches 334; Conservative 1; Mismatches 4;

QY 1053 GGACGAGAAAATCAATTCGTAGTCACTAATGTGTGTTAAACCTGGAGTGAACGACAT 1112
DB 113 GGACGAGAAAATCAATTCGTAGTCACTAATGTGTGTTAAACCTGGAGTGAACGACAT 172

QY 1113 GAATTCCTCGTGGAAACACCTCGCATATGGCGAGTTAAGGATCTGCGAATACCGCCGA 1172
DB 173 GAATTCCTCGTGGAAACACCTCGCATATGGCGAGTTAAGGATCTGCGAATACCGCCGA 232

QY 1173 TCGCATCTGGAGCGGACGCTGCTGATGTACAACAGTCGGATAGGATTTGACGGCAC 1232
DB 233 TCGCATCTGGAGCGGACGCTGCTGATGTACAACAGTCGGATAGGATTTGATGGCAC 292

QY 1233 CTACACAGCAACGCTGGTGGTGGCGAACCAACGGCTCGTGTCTATACGTTCCGCGGGAT 1292
DB 293 CTACACAGCAACGCTGGTGGTGGCGAACCAACGGCTCGTGTCTATACGTTCCGCGGGAT 352

QY 1293 CTTCAAGTCGACGTCGAAGATCGACATCACGTGGTTCCTCTCGATGACCGGTCGCA 1352
DB 353 CTTCAAGTCGACGTCGAAGATCGACATCACGTGGTTCCTCTCGATGACCGGTCGCA 412

QY 1353 GATGAAGTTCGGCAGTTGACCTACGCGGATTCAGCT 1391
DB 413 GATGAAGTTCGGCAGTTGACCTACGCGGATTCAGCT 451

```
Db
425 GATGAAGTTCGGCAGTTGGACCTAGACGCGATTCCAGGT 463

RESULT 4
CNS0006F9
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC:
BACR35F05 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL073676
AL073676.1 GI:4953355
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1007)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1. .1007
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR35F05"
/clone_lib="RPCI-98"
/note="end : TET3"

FEATURES
source
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR35F05"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN
Query Match 14.3%; Score 331.2; DB 29; Length 1007;
Best Local Similarity 98.2%; Pred. No. 1.4e-69;
Matches 333; Conservative 1; Mismatches 5; Indels 0; Gaps 0;

QY 1053 GGACGAGAAAAATCAATTGCTAGTCACTAATGTGTGTTAAACTGGAGTGGAAACGACAT 1112
Db 125 GGACGAGAAAAATCAATTGCTAGTCACTAATGTGTGTTAAACTGGAGTGGAAACGACAT 184
QY 1113 GAATCTCCGCTGGAAACACCTCCGACTATGCGGAGTTAAGGATCTCGGAATACCGCGCA 1172
Db 185 GAATCTCCGCTGGAAACACCTCCGACTATGCGGAGTTAAGGATCTCGGAATACCGCGCA 244
QY 1173 TCGCATCTGGAGCCGGAGCTGTGATGTACAACTGCGGATGAGGGATTTCACGGCAC 1232
Db 245 TCGCATCTGGAGCCGGAGCTGTGATGTACAACTGCGGATGAGGGATTTCACGGCAC 304
QY 1233 CTACACAGCAAGCTGGTGGGAAACACCGCTCGTGTCTATAGTTCCCGCGGGAT 1292
Db 305 CTACACAGCAAGCTGGTGGGAAACACCGCTCGTGTCTATAGTTCCCGCGGGAT 364
QY 1293 CTTCAAGTCGACGTGCAAGATCGACATCACGTGTTCCCTTCGATGACACGCGGTGCGA 1352
Db 365 CTTCAAGTCGACGTGCAAGATCGACATCACGTGTTCCCTTCGATGACACGCGGTGCGA 424
QY 1353 GATGAAGTTCGGCAGTTGGACCTAGACGCGATTCCAGCT 1391
|||||
```

```
Db
425 GATGAAGTTCGGCAGTTGGACCTAGACGCGATTCCAGGT 463

RESULT 4
CNS0006F9
LOCUS
DEFINITION
Drosophila melanogaster genome survey sequence TET3 end of BAC:
BACR35F05 of RPCI-98 library from Drosophila melanogaster (fruit
fly), genomic survey sequence.
AL064281
AL064281.1 GI:4944356
GSS.
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 978)
Genoscope.
Direct Submission
Submitted (02-JUN-1999) Genoscope - Centre National de Sequencage :
BP 191 91006 EVRY cedex - FRANCE (E-mail : segref@genoscope.cns.fr
- Web : www.genoscope.cns.fr)
Determination of this BAC-end sequence was carried out as part of a
collaboration with the Berkeley Drosophila Genome Project (BDGP).
The BDGP is constructing a physical map of the Drosophila
melanogaster genome using these BACs. For further information
please see http://www.fruitfly.org The BDGP Drosophila
melanogaster BAC library was prepared by Kazutoyo Osoegawa and
Aaron Mamoser in Pieter de Jong's laboratory in the Department of
Cancer Genetics at the Roswell Park Cancer Institute in Buffalo,
NY. The library is named RPCI-98 and was constructed by partial
EcoRI digestion of Drosophila DNA provided by the BDGP from the
isogenic strain y2; cn bw sp, the same strain used for the BDGP's
P1 and EST libraries. A more detailed description of the library
and how to order individual BAC clones, the entire library, or
filters for hybridization from the BACPAC Resource Center can be
found at http://bacpac.med.buffalo.edu/drosophila_bac.htm.
Location/Qualifiers
1. .978
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR35F05"
/clone_lib="RPCI-98"
/note="end : TET3"

FEATURES
source
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"
/clone="BACR35F05"
/clone_lib="RPCI-98"
/note="end : TET3"

ORIGIN
Query Match 14.0%; Score 323.8; DB 29; Length 978;
Best Local Similarity 99.1%; Pred. No. 8.9e-68;
Matches 336; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 1053 GGACGAGAAAAATCAATTGCTAGTCACTAATGTGTGTTAAACTGGAGTGGAAACGACAT 1112
Db 117 GGACGAGAAAAATCAATTGCTAGTCACTAATGTGTGTTAAACTGGAGTGGAAACGACAT 175
QY 1113 GAATCTCCGCTGGAAACACCTCCGACTATGCGGAGTTAAGGATCTCGGAATACCGCGCA 1172
Db 176 GAATCTCCGCTGGAAACACCTCCGACTATGCGGAGTTAAGGATCTCGGAATACCGCGCA 235
QY 1173 TCGCATCTGGAGCCGGAGCTGTGATGTACAACTGCGGATGAGGGATTTCACGGCAC 1232
Db 236 TCGCATCTGGAGCCGGAGCTGTGATGTACAACTGCGGATGAGGGATTTCACGGCAC 295
QY 1233 CTACACAGCAAGCTGGTGGGAAACACCGCTCGTGTCTATAGTTCCCGCGGGAT 1292
Db 296 CTACACAGCAAGCTGGTGGGAAACACCGCTCGTGTCTATAGTTCCCGCGGGAT 355
QY 1293 CTTCAAGTCGACGTGCAAGATCGACATCACGTGTTCCCTTCGATGACACGCGGTGCGA 1352
Db 356 CTTCAAGTCGACGTGCAAGATCGACATCACGTGTTCCCTTCGATGACACGCGGTGCGA 415
QY 1353 GATGAAGTTCGGCAGTTGGACCTAGACGCGATTCCAGCT 1391
|||||
```


Db	416	GATGAAGTTCCGCGAGTTGGACCTACGACGATCCAGGT	454
RESULT 5			
AL530299		1201 bp	linear
LOCUS	AL530299	Homosapiens	NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens
DEFINITION			cDNA clone CS0DD007YP05 5-PRIME, mRNA sequence.
ACCESSION	AL530299		
VERSION	AL530299.2		GI:31068132
KEYWORDS			EST.
SOURCE			Homo sapiens (human)
ORGANISM			Homo sapiens
REFERENCE			Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
AUTHORS			Li, W.B., Gruber, C., Jessee, J. and Polayes, D.
TITLE			Full-length cDNA libraries and normalization
JOURNAL			Unpublished (2001)
COMMENT			On Feb 13, 2001 this sequence version replaced gi:12793792.
			Contact: Genoscope
			Genoscope - Centre National de Sequencage
			BP 191 91006 EVRY cedex - France
			Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
			was normalized. Library was constructed by Life Technologies, a
			division of Invitrogen. This sequence belongs to sequence cluster
			7646.r For more information about this cluster, see
			http://www.genoscope.cns.fr/
			cgi-bin/cluster.cgi?seq=CS0DD007CH03QP1&cluster=7646.r. Contact :
			Feng tiang Email : fliang@lifetech.com URL :
			http://fulllength.invitrogen.com/ InvitroGen Corporation 1600
			Faraday Avenue Genoscope sequence ID : CS0DD007CH03QP1.
FEATURES			Location/Qualifiers
			1..1201
			/organism="Homo sapiens"
			/mol_type="mRNA"
			/db_xref="taxon:9606"
			/clone="CS0DD007YP05"
			/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"
			/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"
			/note="1st strand cDNA was primed with a NotI-oligo (dT)
			primer. Five prime end enriched, double-strand cDNA was
			digested with Not I and cloned into the Not I and EcoR V
			sites of the pCMVSPORT 6 vector. Library was normalized."
ORIGIN			
			Query Match 12.2%; Score 281.8; DB 9; Length 1201;
			Best Local Similarity 59.8%; Pred. No. 1.7e-57;
			Matches 503; Conservative 5; Mismatches 322; Indels 11; Gaps 2;
QY	905	CTGCTAAAGTTGGCTAGCAGATATCATGAAAAGACTGTACACGATCTTTGGATC	964
Db	170	CTCTCGCAGTCTCCCTGCAACGAGTTCACAGGAAGCTTTACAGGAGCTGGTCA	229
QY	965	CTTAAATACACTAGACGTCCTCTCAATGAATCGGACCGCTTACAAATTAAGCTTG	1024
Db	230	ACTAAATCCCTTGAGAGGCCGCGGCAATGACTCGCAACACTCACCGTCTACTTCT	289
QY	1025	GTTTAACTTTAATGCAAAATTTATCGATGTGGACGAGAAAATCAATTCGTAGTCATTA	1084
Db	290	CCCTGAGCTCTCGCAGATCATGGACGTGGATGAGAAGAACCAAGTTTAAACCACCA	349
QY	1085	TGTGGTTAAACTGGAGTGAAGACGATCAATTCGCTGGGACACCTCCGACTATGGCG	1144
Db	350	TTTGGCTGCAAAATGCTTGGACAGATCACTATTACAGTGAATGTGTCAGAAATCCAG	409
QY	1145	GAGTTAAGGATCTCGAAATACCGCCATCGCATCTGGAAGCGCGGACGTGCTGATGAC	1204
Db	410	GGGTGAAGACTGTTCGTTTCCAGATGGCCAGATTGGAAACAGACATCTCTCTATA	469
QY	1205	ACAGTCGGATGAGGATTTGACGGCACTTACAGACGAACGTGGTGGTGGCGAACACG	1264

```
/mol_type="mrna"  
/db_xref="taxon:7227"  
/clone="GH15518"  
/sex="male and female"  
/dev_stage="adult"  
/lab_host="DHS - alpha"  
/note="Organ: head; Vector: pOT2; Site 1: EcoRI; Site 2:  
XhoI; Sized fractionated cDNAs were directly ligated into  
pOT2. Plasmid cDNA library."
```

ORIGIN

```
Query Match 11.5%; Score 265; DB 9; Length 607;  
Best Local Similarity 71.4%; Pred. No. 1.7e-53;  
Matches 349; Conservative 0; Mismatches 140; Indels 0; Gaps 0;  
  
Qy 909 TAAAGTTTGGCTACGAGATATCATGAAAGAGACTGTTACAGATCTTTTGGATCCCTTA 968  
Db |||||  
Qy 106 TAAAGAAAGCTGTCAAGGACCTCATGAAAGCGCTGTGAACCATCTGCTGTCCACCTA 165  
Db |||||  
Qy 969 TAATACACTAGAACGTCCTCTCAATGAATCGGACCCGTTACAAATTAAGCTTTGGTTT 1028  
Db |||||  
Qy 166 CAATACGCTGAGGACGCCGTGGCCAAATGAATCGGAGCCCTGGAGGTTAAGTTCCGACT 225  
Db |||||  
Qy 1029 AACTTTAATGCAATTATCGATGTGGACGAGAAATCAATTGCTAGTCACTAATGTGTG 1088  
Db |||||  
Qy 226 GACGCTGCAGCAGATCATCGACGTGGACGAGAAATCAGCTTCTCATACGAATCTTTG 285  
Db |||||  
Qy 1089 GTTAAACTGGAGTGGAAACGATGATCTCCGCTGGAAACACCTCCGACTATGCGGAGT 1148  
Db |||||  
Qy 286 GCTTTCGTTGAGTGGAAACGACTACAATCTCGCTGGAATGAACGGAATACGCGGGGT 345  
Db |||||  
Qy 1149 TAAGGATCTCGAATACCGCGCATCGCATCTGGAAGCGGAGCTGCTGATGTACACAG 1208  
Db |||||  
Qy 346 CARGATCTAGCAATCAGCCCAACAGCTGTGGAAGCCGACGTGCTCATGTACAAACAG 405  
Db |||||  
Qy 1209 TCGGATGAGGATTTGACGGCACCTACAGAACAGCTGTGTGTGGGAAACACGCTC 1268  
Db |||||  
Qy 406 CGCGGATGAGGATTCGATGGCAGTATCACACAGGTTGTGTCAAAATGCGGCGAG 455  
Db |||||  
Qy 1269 GTCTCTATAGTTCGCGCGGGATCTCAAGTGCAGCTGCAAGATCGACATCAGTGGTT 1328  
Db |||||  
Qy 466 TTGTCTGTAGTGTCCCTGATCTTCAAGAGCACATGCAAGATGACATCAGTGGTT 525  
Db |||||  
Qy 1329 CCCCTTCGATGACAGGGTGCAGATGAAGTTCGGAGTGGACCTACAGCGAGATCCA 1388  
Db |||||  
Qy 526 CCATTGTGACCAATGCGGAATGAAATTCGGTAGTTGGACITACCATGGAATCA 585  
Db |||||  
Qy 1389 GCTGGATTT 1397  
Db |||||  
Qy 586 GTTGGATTT 594  
Db |||||
```

RESULT 7

```
CD013901  
LOCUS 2296 bp mrna linear EST 21-OCT-2003  
DEFINITION Single gene library Homo sapiens cDNA, mRNA sequence.  
ACCESSION CD013901  
VERSION CD013901.1 GI:37777431  
KEYWORDS EST.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
1 (bases 1 to 2296)  
Jin, P., Fu, G.K., Wilson, A.D., Yang, J., Chien, D., Hawkins, P.R.,  
Au-Yang, J. and Stuve, L.I.  
PCR isolation and cloning of novel splice variant mRNAs from known  
drug target genes  
Unpublished (2003)  
Contact: Jin, P.  
Incyte Corporation  
3160 Porter Drive, Palo Alto, CA 94304, USA
```

```
Tel: 650 621 8639  
Fax: 650 621 8965  
Email: pjin@incyte.com
```

FEATURES

source

Location/Qualifiers

```
1..2296  
/organism="Homo sapiens"  
/mol_type="mrna"  
/db_xref="taxon:9606"  
/clone_lib="Single gene library"  
/note="Vector: pDrive Cloning Vector; RT-PCR was performed  
using gene-specific primers flanking the open-reading  
frame. PCR products were subcloned into pDrive Cloning  
Vector and sequenced completely using M13 forward and  
reverse primers. Sequencing gaps were closed by  
re-sequencing using primers flanking the gapped areas."
```

ORIGIN

```
Query Match 10.9%; Score 251; DB 14; Length 2296;  
Best Local Similarity 58.1%; Pred. No. 7.2e-50;  
Matches 462; Conservative 0; Mismatches 330; Indels 3; Gaps 1;  
  
Qy 933 TGAAGAGAGACTGTTACACGATCTTTTGGATCCCTTATAATACACTAGAACGTCGCTTCT 992  
Db |||||  
Qy 349 TGAGGACCGGCTCTTCAAACACCTTTCGGGGCTACAACGCTGGGCGGCCCGGTGCC 408  
Db |||||  
Qy 993 CAATGAATCGGACCCGTTACAATTAGCTTTGGTTTAACTTTAATGCAAAATTAATCGATG 1052  
Db |||||  
Qy 409 CAACACTTCAGACGTGGTGATTTGTGCTTTGGACTGTCCATCGCTCAGCTCATCGATGT 468  
Db |||||  
Qy 1053 GGACGAGAAAAATCAATTGCTAGTACTAATGTGTGTTAAACTTAAACTGAGTGGAAACGACAT 1112  
Db |||||  
Qy 469 GGATGAGAGAACCAATGATGACCAACACGCTTGGCTAAACAGAGGTGGAGCGACTA 528  
Db |||||  
Qy 1113 GAATCTCCGCTGGAAACACCTCCGACTATGGCGAGTTAAGGATCTGCGAATACCGCGCA 1172  
Db |||||  
Qy 529 CAAACTGCGCTGGAAACCCCGCTGATTTTGGCAACATCACATCTCTAGGGTCCCTCTGA 588  
Db |||||  
Qy 1173 TCGCATCTGGAAGCGGACGCTGCTGATGTAACAAGTGGCGATGAGGGATTTGACGGCAC 1232  
Db |||||  
Qy 589 GATGATCTGATCCCGACATTTCTCTACAACAAATGACATGGGAGTTTGCAGTGAC 648  
Db |||||  
Qy 1233 CTACACAGAACGCTGTGTGGCGGACACGCTGTGTCTATACGTTCCGCGGGGAT 1292  
Db |||||  
Qy 649 CCATATGACCAAGCCCACTCTTCTCCAGGCACTGTGCACTGGGTGCCCCGGCCAT 708  
Db |||||  
Qy 1293 CTTCAAGTCGACGTGCAAGATCGACATCACGCTGGTTCCCTTCGATGACACGCGTCCGA 1352  
Db |||||  
Qy 709 CTACAAGAGCTCTGCAGCATCGACGTCACTTCTTCCCTTCGACCCAGAGACTGCAA 768  
Db |||||  
Qy 1353 GATGAAGTTCCGAGTGGACCTAGACGATTCAGCGGATTCGAGTGGATTTCAATTAACAAGATGA 1412  
Db |||||  
Qy 769 GATGAAGTTGGCTCTCTGGACTTATGACAAAGGCAAGATCGACCTGGAGCAGATGGAGCA 828  
Db |||||  
Qy 1413 AACTGGCGGTGATATCAGCAGTTTACGCTCTCAACGGCGAGTGGAACTACTGGGTGGCC 1472  
Db |||||  
Qy 829 GACTGT---GGACCTGAGGACTACTTGGGAGAGCGGCGAGTGGGCCATCGTCAATGCCAC 885  
Db |||||  
Qy 1473 CGGCAAACTGAACGAGATCTATTACAACCTGCTGCCCGGAACCCCTATATAGACATCACTT 1532  
Db |||||  
Qy 886 GGGCACTTACAACAGCAAGAAGTACGACTGTCTGCGCGGAGATCTACCCGACGTCACCTA 945  
Db |||||  
Qy 1533 CGGCATCATATCCGCGGACGAACTGATCTATTCTTCTTAAACCTGATCATACCTTGTGT 1592  
Db |||||  
Qy 946 CGCTTTCTGTCATCCGCGGCTGCGCTCTTCTACACCATCAACCTCATATCCCTGCCT 1005  
Db |||||  
Qy 1593 ACTGATTGCTCCATGGCCCTTCTCGGATTCACCTTCCGCGCGAGATTCGGTGGAAAAAT 1652  
Db |||||  
Qy 1006 GTCATCTCTGCTCACTGCTGTCTTCTTACCTGCCCTCCGACTCGGCGGAGAGAT 1065  
Db |||||  
Qy 1653 ATCGCTGGGTGTTACCATCTTCTGCTGACCGGTGTTCTGGAATATGTTGCCGAGAC 1712  
Db |||||  
Qy 1066 CACGCTGTGCTATTCGGTGTCTGCTCACTACCGTCTTCTCTGCTGCTCATCACTGAGAT 1125  
Db |||||
```

```
QY 1713 AATCGCGGTACTTC 1727
Db 1126 CATCCCGTCCACCTC 1140

RESULT 8
BU915857
LOCUS
DEFINITION AGENCOURT 10492745 NICHD_XGC_001 Xenopus laevis cDNA clone
IMAGE:6642638 5', mRNA sequence.
BU915857
ACCESSION BU915857.1 GI:24097771
VERSION
KEYWORDS EST.
SOURCE
ORGANISM
Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus.
REFERENCE
1 (bases 1 to 922)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: ccapbs-remail.nih.gov
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
Cloning Distribution: Agencourt Bioscience Corporation
Cloned through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUM14228 row: f column: 14
High quality sequence stop: 746.
Location/Qualifiers
1. 922
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6642638"
/tissue_type="oocytes"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD_XGC_001"
/note="Vector: pCMV-SF0T6; Site 1: NotI; Site 2: SalI;
Cloned unidirectionally. Primer: Oligo dT. Average insert
size 2.2 kb. Constructed by Life Technologies."
ORIGIN
Query Match 10.2%; Score 235.2; DB 13; Length 922;
Best Local Similarity 60.0%; Pred. No. 3.8e-46;
Matches 415; Conservative 0; Mismatches 268; Indels 9; Gaps 1;
QY 1051 GTGACGAGAAATCAATGTCAGTAAATGTTGGTTAAACTGAGTGGACAC 1110
Db 37 GAGGATGAAAAAATCAAGTACTAACAACAACATATGCTGCTGCAATGGAATGATAGAT 96
QY 1111 ATGATCTCCGTGGAACACCTCCGACTATGCGGAGTTAAGATCTGGAATACCGCG 1170
Db 97 TACTATTTGAGTGAACATGTTGAAATACCTCGGGGTAAATAATGTTGTTTCAAGT 156
QY 1171 CATCCCATCTGGAACCGGACGCTGCTGATGTATCAACAGTCCGGATGAGGATTTGACGGC 1230
Db 157 GGACAGATTTGGAAGCTGATATTCTTCTATACAGTGCAGATGATAGATTGATGCC 216
QY 1231 ACCTACCGACGAAAGTGGTGGGGAACAACGCGCTCGTGTATACGTTCCGCGGGG 1290
Db 217 TCCTTTTCACTAATGTTACTGTGTAATCTCCAGTGACACTGCGCAATATCTGCCACGAGG 276
QY 1291 ATCTTCAAGTCGACGTCGAAGATCGACATCATCGTTGTTCCCTTCGATGACCAAGCGTGC 1350
Db 277 ATTTTCAAGTTCGTTGATACATAGTGTGCAATGTTCCCTTTTGACATACAGAAATG 336
QY 1351 GAGATGAAGTTCGGCAGTTGGACCTACGACGAGTCCAGCTGGATTCAATTACAAGAT 1410

337 AACCTGAAGTTTGGCTCTTGGACCTATGTTGGATGGTCACTTGAACCTGCAGATGCTAGAA 396
1411 GAAACTGGCGGTGATATCAGCAGTTACGTCTCAACGCGAGTGGGAACCTACTCGGTGTG 1470
397 TCA-----GACATATCTGAATATATCGTAATGAGAGTGGGACCTTGTAGCGGTT 447
1471 CCGGGAAAACGTACAGAGATCTATTACAACCTGCTGCCCGGAACCTATATAGACATCACC 1530
448 CTTGAAAAAGAGTGAAGATTACTATGAATGCTGCAAGGAACCTGACCTTGTATGTGACA 507
1531 TTGCCATCATCATCCCGCGACGAACTGTAATTTCTTCAACCTGATCATPACCTTGT 1590
508 TTTACGATAAATAGAGCGCAGGACTCTATACGAGCTTAACCTTTAAATTCCTTGT 567
1591 GTACTGATGCTCCATGCGGCTGCTCGGATTCACCTGCCCGCAGATTCGGGTGAAAAA 1650
568 GTACTGATATCTGCTCTAGCCCTCTCTGGGTTCTCTGCTTCAGCAGACTCAGGAGAGAG 627
1651 TTATCGCTGGGTGTACCATCTTCTCTCGCTGACCGGTCTTCTGAATATCGTTGCCGAG 1710
628 ATCTCACTTGTGTAACAGAGCTTATTTGCCCTCACTGTGTTTCATGCTGTTGGTAGCTGAG 687
1711 ACAATGCGGCTACTTCCGATGCGGTGCCATT 1742
688 ATTATGCCAGCAACCTCAGACTCTGTGCCACT 719

RESULT 9
AK034228
LOCUS
DEFINITION Mus musculus adult male dienecephalon cDNA, RIKEN full-length
enriched library, clone:930165116 product:cholinergic receptor,
nicotinic, alpha polypeptide 4, full insert sequence.
AK034228
ACCESSION AK034228.1 GI:26329798
VERSION
KEYWORDS HTC; CAP trapper.
SOURCE
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 Carninci, P. and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
MEDLINE
PUBMED 10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
MEDLINE
PUBMED 11042159
3
Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Wakahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
MEDLINE
PUBMED 11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
JOURNAL
```


TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
92279253
10349636
2
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsuna, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Ozawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

The PANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3230)

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M.,
Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M.,
Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N.,
Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,
Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akashira, S.,
Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasumishi, A.,
Muramatsu, M. and Hayashizaki, Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama,
Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222,
Fax: 81-45-503-9216]

TITLE
JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL: http://genome.gsc.riken.go.jp/
URL: http://phantom.gsc.riken.go.jp/
Location/Qualifiers

FEATURES
source

1. .3230
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:C630019M18"
/db_xref="MGI:2417536"
/db_xref="taxon:10090"

CDS

/clone="C630019M18"
/sex="male"
/tissue type="hippocampus"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="adult"
36. .1925
/notes="unnamed protein product; cholinergic receptor,
nicotinic, alpha polypeptide 4 (MGD MGI:87888,
GBINM_015730, evidence: BLASTN, 99%, match=1946)
putative"
/codon_start=1
/protein_id="BAC38788.1"
/db_xref="GI:26350297"

/translation="MEIGSGAGPPPLLLPLLLLTGLLPASSHIFTRAHAERLLK
RLFGYNKSRPVANISDVLRVFGSLIAQLIDVDEKNOMTTNNVWQVHDKYKLR
DPGDYENVTIRIPSELIWRPDIIVLYNADGFAVTLTKAHLFYDGRVOWTPAIYK
SSCSIDVTFFPQQCTMKFGSWYDKAKIDLVSMHSRVDQDQFWSGEMVIVDVG
TYNRKTECCAEIYFDITYAFIRRLPLFTINLIIFCLLSCLTIVLVEYLPSSGCK
VTLCISVLLSLTFLLEIIPSTLSVLIPIGILYLLFTWIFVLSIVITVFLVNVHH
RSPRTHTMPAMVRRVLDIVRLLFMKRPVSKDNCRRLLIESMHMANAPRFPPEP
EPGLGDCNQGLSPATFCNMDTAVETQTCRSPSHKVPDLKTSVEKASPCSPG
SCHPPNSSGAPVLIKARSLVQVHPVSSQEAAGSIRCSRISIQYCVSQDGAASITK
PTGSPASLKRPSQLPVSQDOTSPCKCTCKBPSPVPTVLKAGTKAPPHLPSPAL
TRAVEGVQYIADHLKADDTDFSVKDKYVAMVIDRFLMWFIIVCLLTGTLGFLPPW
LAGMI"

ORIGIN

Query Match	10.1%	Score 232.6;	DB 11;	Length 3230;
Best Local Similarity	56.6%	Pred. No. 2.5e-45;		
Matches 452;	Conservative 0;	Mismatches 344;	Indels 3;	Gaps 1;
Qy	929	ATCATGAAAAGAGACTGTTACAGATCTTTTGGATCCCTTATAATACACTAGAACGTCGCG	988	
Db	145	ATGGGAGGAGCGCTCTCTGAAGAGACTTCTCTGGCTACACAAGTGGTCTCGGCAG	204	
Qy	989	TTCTCATGAATCGGACCGCTTCAATTAAGCTTTGTTTAACTTTAATGCAATTAATCG	1048	
Db	205	TAGCCAATATCTAGATGTGGTCTCTTGGCTTTGGCTTGTGCAATGCTCAGCTCAATG	264	
Qy	1049	ATGTGGACGAGAAAAATCAATTGCTAGTCACTAATGTGTGTTTAACTGGAGTGAACG	1108	
Db	265	ATGTGGATGAGAAAAACAGATGATGAGCACCACAGTGTGGTGAACGAGAGTGGCATG	324	
Qy	1109	ACATGAATCTCCGCTGGAACACCTCCGACTATGGCGAGTTAAGATCTCGGAATACGCG	1168	
Db	325	ACTACAACTCGCTGGAGCCCTGCTGACTACGAGAAATGTCACTCCATCCGATCCCAT	384	
Qy	1169	CGCATCGCATCTGGAACCGGACGCTGCTGATGATACACAGTCCGATGAGGATTTGACG	1228	
Db	385	CTGAACCTCATCTGGAGGCTGACATCGTCTCTTACAAACGCGACGGGACITTTGAG	444	
Qy	1229	GCACCTACGACGAAAGCTGGTGGTGGGAAACACGGCTCGTGTCTATACGTTCCGCGCG	1288	
Db	445	TCACCCACCTAACCAAGCCACCTGTTCTATGATGGCGGTGTCAGTGGACACCCCGG	504	
Qy	1289	GGATCTTCAAGTCGACGTCGAAGATGCAATACAGTGGTTCCTCCCTTCGATGACGACGGT	1348	
Db	505	CCAATATAGAGCTCTCTGACAGCTCGACGTCACGCTTCTCCCTTCGACGAGCAACT	564	
Qy	1349	CGGAGATGAAGTTCGGCAGTTGGACCTACGACGATTCACGCTGGATTTACAATTAACAG	1408	
Db	565	GTACCATGAAGTTTGGTCTTGGACCTACGACGAGCCAAAGATTGACTT---GGTGACA	621	
Qy	1409	ATGAACCTGGCGGTGATATCAGCAGTTACGTCCTCAACGCGAGTGGAACTACTGGGTG	1468	
Db	622	TGCACAGCGGTGTGGACCACTGGACTTCTGGGAAAGTGGGAGTGGGTCTATTGGGATG	681	
Qy	1469	TGCCCGGCAAGGTAAACAGATCTATTACAACCTGCTCCCGGACCCCTATATAGACATCA	1528	
Db	682	CCGTGGGCACTTACAAACACGAGAGTATGAATGCTGTGCCGAGATCTATCTCGATCA	741	
Qy	1529	CCTTCGCCATCATCATCCGCGACGAACTGTACTTTTCTTCAACCTGATCATACCTT	1588	

JOURNAL	PUBMED	REFERENCE	AUTHORS	TITLE	JOURNAL	COMMENT	FEATURES	source	gene	ORIGIN
gene trios	Science 302 (5652), 1960-1963 (2003)	14671302	2 (bases 1 to 1436)	Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarin, A., Todd, M.A., Tanenbaum, D.M., Civello, D.R., Lu, F., Murphy, B., Ferrera, S., Wang, G., Zheng, X.H., White, T.J., Smitsky, J.J., Adams, M.D. and Cargill, M.	Direct Submission	Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive, Rockville, MD 20850, USA	This sequence was made by sequencing genomic exons and ordering them based on alignment	Location/Qualifiers		
								1. .1436	/organism="Homo sapiens"	
									/mol_type="genomic DNA"	
									/db_xref="taxon:9606"	
									<1.->1436	
									/gene="CHRNA3"	
									/locus_tag="HCMI369"	
										Query Match 9.8%; Score 225.4; DB 29; Length 1436;
										Best Local Similarity 56.1%; Pred. No. 1.1e-43;
										Matches 446; Conservative 0; Mismatches 346; Indels 3; Gaps 1;
Qy	933	TGAAAAGAGACTGTTACAGCATCTTTTGGATCCTTTATATACACTAGAACGTCCTTCT	992							
Db	20	TGAGCACCGTCTATTGAGCGGTGTTGAAGATTACAATGAGATCATCGGCGCTGTGC	79							
Qy	993	CAATGAATCGACCCCGTTACAATTAAAGCTTTGGTTTAACTTTAATGCAAAATTATCGATGT	1052							
Db	80	CAACGTCTCGACCCAGTCATCATCCATTTCGAGTGTCCATGCTCAGCTGTTGAAGGT	139							
Qy	1053	GGACGAGAAAATCAATTTGCTAGTCACTAATGTGTGGTTAAACTGAGGTGGAACACAT	1112							
Db	140	GGATGAAGTAAACAGATCATGGAGACCAACCTGTGGCTCAAGCAAAATCTGGAATGAC	199							
Qy	1113	GAATCTCCGCTGGAACACCTCCGACTATGGGAGTTAAGGATCTCGAATACCGCGCA	1172							
Db	200	CAAGCTGAAGTGAAACCCCTCTGACTATGTGGGAGAGTTCAAGCGTGTCCCTGCACA	259							
Qy	1173	TCGCATCTGGAACCGGACGTGCTGATGTACAACAGTCGGATGAGGGATTTGACGGCAC	1232							
Db	260	GAAGATCTGGAAGCCAGACATTTGCTGTGTATACAATGCTGTGGGATTTCCAGGTGGA	319							
Qy	1233	CTACGACAGAAAGTGTGTGTGGGAAACACGGTGTGTCTATACGTTCCGCGGGGAT	1292							
Db	320	CGAACAGACCAAAAGCCCTTACTCAAGTACACTGGGAGGTGACTTTGGATCATCTCGG	379							
Qy	1293	CTTCAAGTCGACGTGCAAGATCGACATCAGTGTGTTCCCTTCGATGACACAGCGGTGCGA	1352							
Db	380	CTTTAAGAGCTCTGTAAATCGACGTGACCTACTTCCGTTGATTTACCAAACTGTAC	439							
Qy	1353	GATGAAGTTCCGGAGTTTGGACCTACGACGGATTCACGTGGATTTACAATTACAAGATGA	1412							
Db	440	CATGAAGTTCCGTTCCCTGGTCTACGATAAGGCGAAAATCGAT---CTGGTCTGTATCGG	496							
Qy	1413	AATCGCGGTGATATCAGCAGTTACGTGCTCAACGGGAGTGGGAATCTAGTGGGTGTC	1472							
Db	497	CTCTTTCATGAACCTCAAGACATATTGGGAGAGCGCGGAGTGGGCCATCATCAAGCCCC	556							
Qy	1473	CGGCAAAAGCTAACGAGATCTATTACAACCTGTGCCCGGAAACCTTATATAGACATCACCTT	1532							
Db	557	AGGCTACAAACACGACATCAAGTACAACCTGTGGAGAGGATCTATACCCGACATCACATA	616							
Qy	1533	CGCCATCATATCCCGCAGAAACACTGTACTATTTCTTCAACCTGTATCATACCTTGTGT	1592							
Db	617	CTCGCTGTACATCCGCGCCCTTGTCTTACACCATCAACCTCATCATCCCTCCCTGCT	676							
Qy	1593	ACTGATTCGCTCCATTCGGCTTGTGCTCGGATTCACCTTCGCGCAGATTCGGGTGAAAAAT	1652							

Db	677	GCTCATCTCTTCTCTCACTGCTGCTGCTCTCTTCTACCTGCGCTCCGACTGCGGTGAGAAGGT	7336
Qy	1653	ATCGCTGGGGTGTACCATCTTGTCTCGCTGACCGGTGTTTCTGAAATATGTTGCCGAGAC	1712
Db	737	GACCGTGTGATTTCTGCTCTCTCTCCCTGACGGTGTCTCTCTGCTGATCACTGAGAC	796
Qy	1713	AATGCCGGCTACTTC	1727
Db	797	CATCCCTTCCACCTC	811
RESULT 13			
AY402874			
LOCUS			
DEFINITION	AY402874	1436 bp DNA linear GSS 15-DEC-2003	
ACCESSION	AY402874	Pan troglodytes CHRNA3 gene, VIRTUAL TRANSCRIPT, partial sequence,	
VERSION	AY402874.1	genomic survey sequence.	
KEYWORDS			
SOURCE	AY402874.1	GI:39758857	
ORGANISM		GSS.	
REFERENCE		Pan troglodytes (chimpanzee)	
AUTHORS		Pan troglodytes	
		Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;	
		Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.	
		1. (bases 1 to 1436)	
		Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,	
		Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,	
		Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,	
		Adams,M.D. and Cargill,M.	
TITLE		Inferring nonneutral evolution from human-chimp-mouse orthologous	
JOURNAL		gene trios	
PUBMED		Science 302 (5652), 1960-1963 (2003)	
REFERENCE		14671302	
AUTHORS		2. (bases 1 to 1436)	
		Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,	
		Todd,M.A., Tanenbaum,D.M., Civeillo,D.R., Lu,F., Murphy,B.,	
		Ferrera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,	
		Adams,M.D. and Cargill,M.	
TITLE		Direct Submission	
JOURNAL		Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,	
COMMENT		Rockville, MD 20850, USA	
		This sequence was made by sequencing genomic exons and ordering	
		them based on alignment.	
FEATURES		Location/Qualifiers	
source		1..1436	
		/organism="Pan troglodytes"	
		/mol_type="genomic DNA"	
		/db_xref="taxon:9598"	
gene		<1..>1436	
		/gene="CHRNA3"	
		/locus_tag="HCM1369"	
ORIGIN			
Query Match		9.6%; Score 222.2; DB 29; Length 1436;	
Best Local Similarity		54.7%; Pred. No. 6.7e-43;	
Matches 435; Conservative 0; Mismatches 357; Indels 3; Gaps 1;			
Qy	933	TGAAAGAGACTGTTACAGCATCTTTTGGATCCCTTATAATACACTAGAACGTCCTCGTTCT	992
Db	20	TGAGCAGCGTCTATTGAGCGGCTGTTTGAGATTACATGAGATCATCGGCCCTGTGGC	79
Qy	993	CAATGAATCGGACCCGGTTACAATTAAGCTTTGGTTTAACTTTAATGCAAAATTAFCGATGT	1052
Db	80	CAACGTGTCTGACCCAGTCATCATCCATTNNNGGTGTCCATGCTNNNTGGTGAAGGT	139
Qy	1053	GGACAGAGAAATCAATGCTAGTCACTAATGTGTGGTTAAACTGGAGTGGACGACAT	1112
Db	140	GGATGAAGTAAACAGATCATGTGAGAGACCAACNTNNGCTCAAGCAAAATCTGGAATGACTA	199
Qy	1113	GAATCTCGCTGGAAACACTCCGACTATGGCGGAGTTAAGGATCTGGGAATACCGCCGCA	1172
Db	200	CAAGCTGAAGTGGAAACCCCTCTGACTATGTTGGGCGACAGTTTCATGCGTGTCCCTGCNN	259
Qy	1173	TGCAATCTGGAAGCGGACGTGCTGATGTAACAACAGTCGGATGAGGATTTGACGGCAC	1232

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:09:28 ; Search time 66.2366 Seconds
(without alignments)
3284.615 Million cell updates/sec

Title: US-09-303-232-2

Perfect score: 4043

Sequence: 1 MKNAQLKTEVDDELWLAV.....MFAIIATIVLLSAPHIIVS 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

A_Geneseq_29Jan04:*

1: Geneseq1980s:*

2: Geneseq1990s:*

3: Geneseq2000s:*

4: Geneseq2001s:*

5: Geneseq2002s:*

6: Geneseq2003as:*

7: Geneseq2003bs:*

8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4043	100.0	770	3	AA50814 D. melano
2	1850	45.8	496	3	AA50815 H. viresc
3	1609	39.8	501	3	AA50816 H. viresc
4	1498.5	37.1	327	4	ABB70382 Drosophil
5	1407.5	34.8	311	4	ABB63683 Drosophil
6	1179.5	29.2	502	2	AAW12368 Neuronal
7	1173.5	29.0	502	7	AADE57308 Rat Prote
8	1173.5	29.0	502	7	ADDA47049 Rat Prote
9	1168	28.9	511	2	AAW12369 Neuronal
10	1159.5	28.7	502	2	AAW44153 Human neu
11	1159.5	28.7	502	2	AAW09025 Neuronal
12	1159.5	28.7	502	3	AAAB24088 Human PRO
13	1159.5	28.7	502	3	AAAB24088 Human PRO
14	1159.5	28.7	502	4	AAAB2690 Nicotinic
15	1159.5	28.7	502	4	AAAB50012 Wild-type
16	1159.5	28.7	502	5	ABB82435 Human neu
17	1159.5	28.7	502	7	ADDA10874 Human neu
18	1159.5	28.7	502	7	ADDA47051 Human PRO
19	1159.5	28.7	502	7	AADE57310 Human PRO
20	1159.5	28.6	502	2	AAW69216 V2747 var
21	1153.5	28.5	502	4	AAAB50015 Mutant hu
22	1149.5	28.4	502	4	AAAB50016 Mutant hu
23	1143.5	28.3	502	4	AAAB50017 Mutant hu
24	1125	27.8	498	4	ABB60432 Drosophil
25	989	24.5	193	4	ABB60716 Drosophil

ALIGNMENTS

RESULT 1

AA50814

ID AAY50814 standard; protein; 770 AA.

XX AC AAY50814;

XX AC

DT 17-FEB-2000 (first entry)

XX DE

XX D. melanogaster acetyl-choline receptor protein from clone Da7.

KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
KW neurotransmission; plant protection agent; conductance; AChR.

XX OS

XX Drosophila melanogaster.

XX DE19819829-A1.

XX PD 11-NOV-1999.

XX PF 04-MAY-1998; 98DE-01019829.

XX PR 04-MAY-1998; 98DE-01019829.

XX (FARB) BAYER AG.

XX Adamczewski M, Oellers N, Schulte T;

XX WPI; 2000-014207/02.

XX N-PSDB; AAZ24475.

XX New nucleic acid encoding a nicotinic acetylcholine receptor from

XX insects, used to identify potential insecticides.

XX Example 1a; Page 12-14; 26pp; German.

XX This invention describes a novel nucleic acid (NA) encoding a nicotinic

XX acetyl-choline receptor (I) from insects which can be used as an

XX insecticide. Inhibitors of (I) interfere with neurotransmission. (I)

XX (also vectors containing it, its regulatory regions, and antibodies

XX directed against (I)-encoded proteins) are used to screen for: (a) plant

XX protection agents that alter conductance of AChR, potentially useful as

XX insecticides, or (b) genes which encode polypeptides that are involved in

XX formation of functionally related AChR in insects. (I) are also used to

XX isolate and characterize the specified regulatory regions and for

XX recombinant production of (II). This sequence represents an acetyl-

XX choline receptor isolated from Drosophila melanogaster

XX Sequence 770 AA;

SQ

Query Match 100.0%; Score 4043; DB 3; Length 770;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 770; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 MKNAQLKLTVEVDDDELAVRLAHCSNFSSSSSTRTSSNQRHNOQLTLQPRSLSTKH 60
 Db 1 MKNAQLKLTVEVDDDELAVRLAHCSNFSSSSSTRTSSNQRHNOQLTLQPRSLSTKH 60

Qy 61 HSNITASECHNSQOQEPASKEDVANHGSRNDQOHLQOLDSSNMLSPKTAATAAGDEA 120
 Db 61 HSNITASECHNSQOQEPASKEDVANHGSRNDQOHLQOLDSSNMLSPKTAATAAGDEA 120

Qy 121 TQOQTNIRLCARQRRLRRRKRKPAENETDIKKQQQLSMPPEKTKRSTDTTSTAAT 180
 Db 121 TQOQTNIRLCARQRRLRRRKRKPAENETDIKKQQQLSMPPEKTKRSTDTTSTAAT 180

Qy 181 TSCPTATMOCRASDNFSPISRHDRVSTATFAWLHVQLVLSLQWQLHVQQRSVL 240
 Db 181 TSCPTATMOCRASDNFSPISRHDRVSTATFAWLHVQLVLSLQWQLHVQQRSVL 240

Qy 241 LFRRIASTAFISYLGSAQLKNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 300
 Db 241 LFRRIASTAFISYLGSAQLKNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 300

Qy 301 LSAKVCLAGYHEKRLHLLDLPYNTLBRPVLNEDPLQSLFGLTLMQIIDVDEKNQLLVT 360
 Db 301 LSAKVCLAGYHEKRLHLLDLPYNTLBRPVLNEDPLQSLFGLTLMQIIDVDEKNQLLVT 360

Qy 361 NVWLKLEWDMNLRWNTSDYGVGKDLRIPPHRIWKPVDVLYNSADEGFGTYQTNVVVRN 420
 Db 361 NVWLKLEWDMNLRWNTSDYGVGKDLRIPPHRIWKPVDVLYNSADEGFGTYQTNVVVRN 420

Qy 421 NGSCLYVPPGIFKSTCKIDITWFFPDQORCEMKGSTYDGFQDLQLODETGGDISYV 480
 Db 421 NGSCLYVPPGIFKSTCKIDITWFFPDQORCEMKGSTYDGFQDLQLODETGGDISYV 480

Qy 481 LNWGWEVLGVPKNEIYVNCPEYIDITFAIIIRRTLYFFNLIIPCVLIASMALLG 540
 Db 481 LNWGWEVLGVPKNEIYVNCPEYIDITFAIIIRRTLYFFNLIIPCVLIASMALLG 540

Qy 541 FTLPDPSGEKLSLGVTLILLSLTVFLNMVAETMPATSDAVPLWIRVFLCWLPIRLMSRP 600
 Db 541 FTLPDPSGEKLSLGVTLILLSLTVFLNMVAETMPATSDAVPLWIRVFLCWLPIRLMSRP 600

Qy 601 GRPILFPTTSCDTSERKHQILSDVELKERSKSLANVLIDDDDFRNCRPMTPGG 660
 Db 601 GRPILFPTTSCDTSERKHQILSDVELKERSKSLANVLIDDDDFRNCRPMTPGG 660

Qy 661 TLPNPAFYRTVYGGDDSGIPGSTRMPDAVTHHTCIKSTVEYELGLILKEIRFTDQ 720
 Db 661 TLPNPAFYRTVYGGDDSGIPGSTRMPDAVTHHTCIKSTVEYELGLILKEIRFTDQ 720

Qy 721 LRKDECDNDIANDKFAAMVVDRLCLIIFTFMAILATIAVLLSAPHIIVS 770
 Db 721 LRKDECDNDIANDKFAAMVVDRLCLIIFTFMAILATIAVLLSAPHIIVS 770

RESULT 2
 ID AAY50815
 AC AAY50815; standard; protein; 496 AA.
 XX
 XX
 DT 17-FEB-2000 (first entry)
 DE
 DE H. virescens acetyl-choline receptor protein from clone Hva7-1.
 KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
 KW neurotransmission; plant protection agent; conductance; AChR.
 XX
 OS Heliothis virescens.
 XX

PN DE19819829-A1.
 XX 11-NOV-1999.
 XX 04-MAY-1998; 98DE-01019829.
 XX 04-MAY-1998; 98DE-01019829.
 XX (PARB) BAYER AG.
 XX Adamczewski M, Oellers N, Schulte T;
 XX WPI; 2000-014207/02.
 XX N-PSDB; AAZ24476.
 XX New nucleic acid encoding a nicotinic acetylcholine receptor from
 XX insects, used to identify potential insecticides.
 XX Example 1a; Page 17-19; 26pp; German.
 CC This invention describes a novel nucleic acid (NA) encoding a nicotinic
 CC acetyl-choline receptor (I) from insects which can be used as an
 CC insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
 CC (also vectors containing it, its regulatory regions, and antibodies
 CC directed against (I)-encoded proteins) are used to screen for: (a) plant
 CC protection agents that alter conductance of AChR, potentially useful as
 CC insecticides, or (b) genes which encode polypeptides that are involved in
 CC formation of functionally related AChR in insects. (I) are also used to
 CC isolate and characterize the specified regulatory regions and for
 CC recombinant production of (II). This sequence represents an acetyl-
 CC choline receptor isolated from *Heliothis virescens*
 XX Sequence 496 AA;
 SQ

Query Match 45.8%; Score 1850; DB 3; Length 496;
 Best Local Similarity 70.0%; Pred. No. 7.2e-161;
 Matches 361; Conservative 30; Mismatches 49; Indels 76; Gaps 7;

Qy 295 LLIYVNLISAKVCLAGYHEKRLHLLDLPYNTLBRPVLNEDPLQSLFGLTLMQIIDVDEK 354
 Db 17 LLLCULWPGARGCYHEKRLHLLDLPYNTLBRPVLNEDPLQSLFGLTLMQIIDVDEK 76

Qy 355 NQLIVTNNWLKLEWDMNLRWNTSDYGVGKDLRIPPHRIWKPVDVLYNSADEGFGTYQ 414
 Db 77 NQLIITNWLKLEWDMNLRWNTSDYGVGKDLRIPPHRIWKPVDVLYNSADEGFGTYQ 136

Qy 415 NVVVRNNGSCLYVPPGIFKSTCKIDITWFFPDQORCEMKGSTYDGFQDLQLODETGG 474
 Db 137 NVVVRNNGSCLYVPPGIFKSTCKIDITWFFPDQORCEMKGSTYDGFQDLQLODETGG 196

Qy 475 DISSVVLNGEWELGVPKNEIYVNCPEYIDITFAIIIRRTLYFFNLIIPCVLIA 534
 Db 197 DISSVVLNGEWELGVPKNEIYVNCPEYIDITFAVIRKRLTYFFNLIIPCVLIA 256

Qy 535 SMALLGFTLPDPSGEKLSLGVTLILLSLTVFLNMVAETMPATSDAVPL----- 581
 Db 257 SMALLGFTLPDPSGEKLSLGVTLILLSLTVFLNMVAETMPATSDAVPLGTYFNCIMFVA 316

Qy 582 -----WIRVFLCWLPIRLMSRP---RPLIERPTTCS 614
 Db 317 SSVVSTTILNLYNHRHADTHMSDWIRCVFLWFLWLRMSRPGSATPPPARVPPPP-- 374

Qy 615 DTSSERKHQILSDVELKERSKSLANVLIDDDDFRNCRPMTPGGTLPNPAFYRTVY 674
 Db 375 -----DLELRERSKSLANVLIDDDFRH-----PQAQPPQCCRYR---- 412

Qy 675 QGDGSGIPGSTRMPDAVTHHTCIKSTVEYELGLILKEIRFTDQLRKDECDNDIANDW 734
 Db 413 GGEENGAG-----LAAHSCF--GVDYELSLILKEIRVITDQMKRDEDEDADISRDW 460

Qy 735 KFAAMVVDRLCLIIFTFMAILATIAVLLSAPHIIVS 770
 Db 461 KFAAMVVDRLCLIIFTFMAILATIAVLLSAPHIIVS 496

sd
1;

QY 1 MKNAQLKLTVDDELWLAHLACSNFSSSSSTRTTSSNQRHNOQLTTILOPESLSTKH 60
 Db 1 MKNAQLKLTVDDELWLAHLACSNFSSSSSTRTTSSNQRHNOQLTTILOPESLSTKH 60
 QY 61 HSNIASQHNSSQOEPAKDEEDVANHGSRNDQOHLQQLDSSNNMLSPKTAATAAGDEA 120
 Db 61 HSNIASQHNSSQOEPAKDEEDVANHGSRNDQOHLQQLDSSNNMLSPKTAATAAGDEA 120
 QY 121 TQOPTNIRLCARQRLRRRRKPKATNETDIKKOQQLSMPFPKTRKSTDTYSTPAAT 180
 Db 121 TQOPTNIRLCARQRLRRRRKPKATNETDIKKOQQLSMPFPKTRKSTDTYSTPAAT 180
 QY 181 TSCTATYMOCRASDNFSPISRHDRVSTATFAWLHVQLVLSLQOQLHVQOQSVL 240
 Db 181 TSCTATYMOCRASDNFSPISRHDRVSTATFAWLHVQLVLSLQOQLHVQOQSVL 240
 QY 241 LFRRIAASTIAFISYLGFAAQLKNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 299
 Db 241 LFRRIAASTIAFISYLGFAAQLKNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 300
 QY 300 NLSAK 304
 Db 301 NLSAK 305

RESULT 5
 ABB63683
 ID ABB63683 standard; protein; 311 AA.
 AC ABB63683;
 XX
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 17841.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 FN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL07786.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from Drosophila and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 17841; 21pp + Sequence Listing; English.
 XX
 CC The invention relates to an isolated nucleic acid detection reagent
 CC capable of detecting 1000 or more genes from Drosophila. The invention is
 CC useful in developmental biology and in elucidating cell signalling and
 CC cell-cell interactions in higher eukaryotes for the development of
 CC insecticides, therapeutics and pharmaceutical drugs. The invention
 CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 CC sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 CC ABB72072). The sequence data for this patent did not form part of the
 CC printed specification, but was obtained in electronic format directly
 CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
 XX

SQ Sequence 311 AA;
 Query Match 34.8%; Score 1407.5; DB 4; Length 311;
 Best Local Similarity 88.6%; Pred. No. 1.8e-120;
 Matches 257; Conservative 18; Mismatches 12; Indels 3; Gaps 1;
 QY 295 LLIV--LNLAKVCLAGYHEKRLHLLDLPYNTLFRPVLNESDPLQLSFGTLTLMQIIDV 351
 Db 19 MLVYGLLLTMTFACAAGPHEKRLHLLDLPYNTLFRPVLNESDPLQLSFGTLTLMQIIDV 78
 QY 352 DEKNQLLVTVWMLKLEWMDNMLRWNTSDYGVKDLRTPPHRIKPDVLMVNSADEGFDGT 411
 Db 79 DEKNQLLVTVWMLKLEWMDNMLRWNTSDYGVKDLRTPPHRIKPDVLMVNSADEGFDGT 138
 QY 412 YQTNVVRNNGSCLYVPPGIFKSTCKIDITWFPDDORCEMKFGSWTYDGFQDLQLODE 471
 Db 139 YATNVVVRNNGSCLYVPPGIFKSTCKIDITWFPDDORCEMKFGSWTYDGFQDLQLODE 198
 QY 472 TGGDISSYVLNGEWELLGVFGKNEIYVNCPCPEYIDITFAIIRRTLYFFENLIPCV 531
 Db 199 AGGDISSFITNGEWDLLGVFGKNEIYVNCPCPEYIDITFAIIRRTLYFFENLIPCV 258
 QY 532 LIASMAILGFTLPDPSGEKLSLGVITILLSTVFLNMVAETMPATSDAVPL 581
 Db 259 LIASMAILGFTLPDPSGEKLSLGVITILLSTVFLNMVAETMPATSDAVPL 308
 RESULT 6
 AAW12368
 ID AAW12368 standard; protein; 502 AA.
 AC AAW12368;
 XX
 DT 17-JUN-1997 (first entry)
 XX
 DE Neuronal alpha-bungarotoxin binding protein alpha subunit.
 KW Neuronal alpha-bungarotoxin binding protein alpha 1; cholinergic;
 KW ligand binding; ion channel.
 XX
 OS Gallus sp.
 XX
 FH Key Location/Qualifiers
 FT Peptide 1..22
 FT Protein /label= Sig_peptide
 FT Protein 23..502
 FT Protein /label= Mat_protein
 XX
 PN US5599709-A.
 XX
 PD 04-FEB-1997.
 XX
 PF 28-SEP-1989; 89US-00413947.
 XX
 PR 28-SEP-1989; 89US-00413947.
 XX
 PA (SALK) SALK INST BIOLOGICAL STUDIES.
 XX
 PI Lindstrom JM, Schoepfer RD;
 XX
 DR WPI; 1997-118297/11.
 DR N-PSDB; AAT59196.
 XX
 CC New isolated neuronal alpha-bungarotoxin-binding protein DNA - used to
 CC screen cholinergic agents and other drugs which may affect ligand
 CC binding, ion channel or other activities of the protein.
 XX
 PS Example; Fig 2A-B; 18pp; English.
 CC The alpha subunit (AAW12368) and alpha2 subunit (AAW12369) amino acid
 CC sequences of chick neuronal alpha-bungarotoxin binding protein (ABBP)
 CC were deduced from newly isolated DNA molecules (AAT59196-97) obcd. from
 CC an 18-day chick embryo cdna library. ABBP subunits can be produced in


```
PN WO200053755-A2.
XX
XX 14-SEP-2000.
XX
XX 06-JAN-2000; 2000WO-US000376.
XX
XX 08-MAR-1999; 99WO-US005028.
XX PR 02-JUN-1999; 99WO-US012252.
XX PR 23-JUN-1999; 99US-0141037P.
XX PR 07-JUL-1999; 99US-0143048P.
XX PR 26-JUL-1999; 99US-0145698P.
XX PR 30-NOV-1999; 99WO-US028313.
XX PR 20-DEC-1999; 99WO-US030911.
XX PR 05-JAN-2000; 2000WO-US000219.
XX
XX (GETH ) GENENTECH INC.
XX
XX Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillan KJ, Roy MA;
XX Watanabe CK, Wood WI;
XX WPI; 2000-572270/53.
XX DR N-PSDB; AAC58395.
XX
XX Thirty PRO polynucleotides encoding PRO polypeptides, useful in the
XX treatment, diagnosis and prevention of cancer.
XX
XX Claim 61; Fig 58; 286pp; English.
XX
XX The present invention describes an isolated antibody that binds to one of
XX the human PRO proteins designated PRO212, PRO290, PRO341, PRO535, PRO619,
XX PRO717, PRO809, PRO830, PRO843, PRO1005, PRO1009, PRO1025,
XX PRO1030, PRO1097, PRO1107, PRO1111, PRO1153, PRO1182, PRO1184, PRO1187,
XX PRO1281, PRO23, PRO39, PRO834, PRO1317, PRO1710, PRO2094, PRO2145 OR
XX PRO2198. PRO antagonists can be used to inhibit tumour cell growth. The
XX PRO polypeptides and nucleotides are useful in the treatment, diagnosis
XX and prevention of cancer. The antibodies and other anti-tumour compounds
XX maybe used to treat various conditions, including those characterised by
XX overexpression and/or activation of the amplified PRO genes. Exemplary
XX conditions or disorders to be treated with such antibodies and other
XX compounds include benign or malignant tumours (e.g., renal, liver,
XX kidney, bladder, breast, gastric, ovarian, colorectal, prostate,
XX pancreatic, lung, vulva, thyroid, hepatic carcinomas, sarcomas,
XX glioblastomas, and various head and neck tumours), leukaemias and
XX lymphoid malignancies, other disorders such as neuronal, glial,
XX astrocytal, hypothalamic and other glandular, macrophagal, epithelial,
XX stromal and blastocoeic disorders, and inflammatory, angiogenic and
XX immunologic disorders. AAC58242 to AAC58366 represent PCR primers and
XX hybridisation probes used in the isolation of the human PRO sequences.
XX AAC58367 to AAC58396 and AAB24057 to AAB24089 represent human PRO
XX polynucleotide and protein sequences given in the exemplification of the
XX present invention
XX
XX Sequence 502 AA;
XX
XX Query Match 28.7%; Score 1159.5; DB 3; Length 502;
XX Best Local Similarity 45.1%; Pred. No. 2.6e-97;
XX Matches 237; Conservative 82; Mismatches 115; Indels 91; Gaps 12;
XX
XX 297 IYNLSAK--VCLAGVHEKRLHDLDPYNTLPERPVNLSDELQLSFGTLMQIDVDE 353
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 8 VWLAAALHVSLSQGFQKLYKELVKNVPLERPVANDSQPLTVVYFSLSLQIMDVDE 67
XX
XX 354 KNOLLVTNVLKLEWDMNLRWNTSDYGVKDLRIPHRITWKFDVLNYSADSGFDGTQY 413
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 68 KNOVLTTNVLQMSWTDHYLQWNVSEYPGVKTRFPDQGIWKFDILLYNSADERFDATFH 127
XX
XX 414 TNNVRRNNGSLVPPGIFKSTCKIDITWPFDDQRCMKFGSWTYDGFOLDLQLODETG 473
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 128 TNNLVNSGHCQYLPQGIKSSCYIDVWFDFVQHKLFKFGWSYGSGLDQMQE--- 184
XX
XX 474 GDTSYVLNNGEWELLGVFGKRNIIYNCPEPYIDITFAIIIRRTLYYFNNLIIPCVLI 533
XX : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
XX 185 ADISGVIPNGEWDLVGIPGKRSEFYECCKEYPDVTFVTMRRRTLYYGLNLLIPCVLI 244
XX
```

```
QY 534 ASMAILGFTLPPDSGEKLSLGVTTLLSLTVFLNVAETMPATSDAVPL----- 581
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
245 SALALLVFLLPADSGEKISLGIITVLLSLTVFMLLVAETIMPATSDSVPLIAQYFASTMIIV 304
QY 582 -----WLRIVFLCWLPLILMSREG-----RLILLEFFPTP 612
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
305 GLSVVTVTVLVQYHHDDPGGKMPKTRVILLNCAWFLRMKRPGEKVRP-ACQHKORR 363
QY 613 CSDTSSERKQILSDVELKERSKSLIANVLIDDDDFRH-----NCRPMTPGGTLPHNPAF 668
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
364 CSLASVE-----MSAVAPPASNGNLIY-----IGFRLDGVHCVP-TP----- 401
QY 669 YRTVYGGDDSGSIGPIGSTRMPDAVTH-----HTCIKSTSEYELGLILKEIRFTDQLRK 723
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
402 -----DSGVV-----CGRMACSPTHDEHLHGQPPGPDPLAKILLEEVRYIANRPRC 449
QY 724 DDECNDIANDKFAAMVVDRLCLIFTFMEAILATIALLSAPHII 768
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
450 QDESEAVCSEWKAACVVDRLCLMAFSVFTIICITGILMSAPNFV 494

RESULT 13
AAB82690
ID AAB82690 standard; protein; 502 AA.
XX
XX AC AAB82690;
XX
XX 15-OCT-2001 (first entry)
XX DT Nicotinic acetylcholine receptor alpha7.
XX DE
XX KW Nicotinic acetylcholine receptor; nAChR; human;
XX KW acetylcholine binding protein; AChBP; mollusc; ligand-binding protein;
XX KW ligand-gated ion channel; crystal; drug design; protein co-ordinate data;
XX KW schizophrenia; Alzheimer's disease; nicotine addiction;
XX KW Tourette's syndrome; therapy; nootropic; neuroprotective.
XX
XX OS Homo sapiens.
XX
XX Key Location/Qualifiers
XX Region 108..115
XX /note= "conserved ligand-binding region, residues Trp108
XX and Tyr115 are essential"
XX Region 171..173
XX /note= "conserved ligand-binding region, residues Trp171
XX and Tyr173 are essential"
XX Region 210..217
XX /note= "conserved ligand-binding region, residues Tyr210,
XX Cys212, Cys213 and Tyr217 are essential"
XX
XX WO200158951-A2.
XX
XX 16-AUG-2001.
XX
XX 09-FEB-2001; 2001WO-EP001457.
XX
XX 10-FEB-2000; 2000EP-00200443.
XX PR 31-OCT-2000; 2000EP-00203810.
XX
XX (TEWE-) STICHTING TECH WETENSCHAPPEN.
XX
XX Smit AB, Sixma TK;
XX WPI; 2001-497071/54.
XX
XX Water-soluble ligand-binding proteins derived from molluscs and analogs
XX of ligand-gated ion channels, useful in drug screening assay, where the
XX drugs identified can be used in the treatment of Alzheimer's disease or
XX schizophrenia.
XX
XX Disclosure; Page 252-254; 260pp; English.
XX
```

CC The present sequence is that of the alpha subunit of human nicotinic
CC acetylcholine receptor (nAChR). The sequence includes regions that are
CC conserved throughout the various nAChR alpha subunits and which are
CC essential for ligand binding. The invention relates to water-soluble
CC ligand-binding proteins derived from molluscs, especially acetylcholine-
CC binding proteins (AChBPs) and analogues of ligand-gated ion channels,
CC their crystals, and their use for screening ligands of ligand-gated ion
CC channels. The water-soluble ligand-binding proteins are capable of
CC forming multimers and are amenable to crystallization. The crystal
CC structure of AChBP is provided, and can be used to generate 3D models of
CC the extracellular ligand-binding domain of ligand-gated ion channels and
CC thus for screening of drugs that act on these ion channels. Chimeric
CC proteins are provided that are capable of binding a ligand of a ligand-
CC gated receptor, and comprise at least the amino acids of the AChBP
CC determining solubility of the AChBP, in the same positions as in the
CC AChBP, and also comprising amino acids determining binding to the ligand.
CC In the chimeric proteins, at least the essential amino acids of at least
CC 1 of the conserved regions of an nAChR have been substituted for the
CC corresponding amino acids, and preferably entire stretches have been
CC substituted. New drugs can be developed that selectively intervene in
CC neuronal signalling pathways, especially where the ligand-gated ion
CC channel is the nAChR, and the related disorder is Tourette's syndrome,
CC Alzheimer's disease, addiction to nicotine or schizophrenia
XX

SQ Sequence 502 AA;
Query Match 28.7%; Score 1159.5; DB 4; Length 502;
Best Local Similarity 45.1%; Pred. No. 2.6e-97;
Matches 237; Conservative 82; Mismatches 115; Indels 91; Gaps 12;
QY 297 IYLNLSAK---VCLAGYHEKRLHLLDPYNTLPRVNLSPDQLSFLGLTMIQIDVE 353
DB 8 VWLALASLLHVSLOGEFQKLYKELVKNYPLERPVANDSQPLTVYFSLILQIMDVE 67
QY 354 KNQLLVTVNWLKLEWDMNLWNTSDYGVKDLRIPPHRIWKPDPVLMVNSADEGPDGTQY 413
DB 68 KNQVLTVNWLQMSWTDHYLQMNVSYPGVKTVRFPDQIWKPDILLVNSADERFDATFH 127
QY 414 TNVVRNNGSLYVPPGIFKSTCKIDITWFFDDQRCMKFGSWTYDGFQDLQLOQDETG 473
DB 128 TNVLNNSGHCQYLPFGIFKSCYDVRWFFDVQHCXKLFKGSWSYGSWLDLQMQE--- 184
QY 474 GDISSYVLNGEWELLGVPGKENEIYNNCCPEYIDITFAIIRRTLYFFNLIIPCVLI 533
DB 185 ADISGYIPNGEWDLVGIPGKRSERFECCKEYPDVTFTVMRRRTLYYGLNLIIPCVLI 244
QY 534 ASMLLGLTLPDSEKLSLGVTLISLTVFLNVAETMPATSDAVL----- 581
DB 245 SALALLVFLPADSGEKISLGITVLLSLTVFMLLVAEIMPATSDVPLIAQYFASTMIIV 304
QY 582 -----WIRIVLCWLPWILRMSRPG-----RPLILEFFPTTP 612
DB 305 GLSVVTVTVIVLYQHHDDPGCKMPKWTETVILLNWCWFLMKRPGEDKVRP-ACQHKQR 363
QY 613 CSDTSSERKHQILSDVELKERSKSLANVLIDDDFRH-----NCRPMTPGGTLPHNPAF 668
DB 364 CSLASVE-----MSAVAPPPASNGNLLY-----IGFGLDGVHCV-TP----- 401
QY 669 YRTVYGGDDGSGIGPIGSTRMPDAVTH-----HTCIKSTVEYELGLILKEIRFTDQLRK 723
DB 402 -----DSGVV-----CGRWACSGPTHDEHLHGGQPPGDPDLAKILEEVRYIANRRC 449
QY 724 DDECDIANDKFAAMVVDRLCLIFFTWFAILATIAVLLSAPHII 768
DB 450 QDESEAVCSEKFAACVVDRLCLMAFSVFTIICITIGILMSAPNFV 494
RESULT 14
ID AAB50012
XX AAB50012 standard; protein; 502 AA.
AC AAB50012;
XX

DT 14-MAR-2001 (first entry)
XX
DE Wild-type human alpha7 ligand gated ion channel.
XX
KW Human; alpha7 nicotinic acetylcholine gated ion channel;
XX 5-hydroxytryptamine; 5-HT3; calcium ion conductance.
OS Homo sapiens.
XX
PN WO2000073431-A2.
XX
PD 07-DEC-2000.
XX
PF 25-MAY-2000; 2000WO-US011862.
XX
PR 27-MAY-1999; 99US-0136174P.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Groppi VE, Wolfe ML, Berkenpas MB;
XX WPI; 2001-061524/07.
DR N-PSDB; AAC90380.
XX
PT Special cell culture medium for treating cells and for inducing mammalian
PT cell lines to conduct calcium ions, comprising specified concentrations
PT of ions of sodium, calcium and potassium at specified pH.
XX
PS Disclosure; Page 61-63; 77pp; English.
XX
CC The present sequence is wild-type human alpha7 nicotinic acetylcholine
CC gated ion channel. The human alpha7 ion channel was used in the
CC construction of an alpha7/5-hydroxytryptamine (5-HT3) chimeric
CC gated ion channel (see AAC90382 and AAB50014). The alpha7/5-HT3 chimeric
CC ion channel can be expressed by recombinant cells in the present
CC invention, resulting in preferential calcium ion conductance by the cells
XX
SQ Sequence 502 AA;
Query Match 28.7%; Score 1159.5; DB 4; Length 502;
Best Local Similarity 45.1%; Pred. No. 2.6e-97;
Matches 237; Conservative 82; Mismatches 115; Indels 91; Gaps 12;
QY 297 IYLNLSAK---VCLAGYHEKRLHLLDPYNTLPRVNLSPDQLSFLGLTMIQIDVE 353
DB 8 VWLALASLLHVSLOGEFQKLYKELVKNYPLERPVANDSQPLTVYFSLILQIMDVE 67
QY 354 KNQLLVTVNWLKLEWDMNLWNTSDYGVKDLRIPPHRIWKPDPVLMVNSADEGPDGTQY 413
DB 68 KNQVLTVNWLQMSWTDHYLQMNVSYPGVKTVRFPDQIWKPDILLVNSADERFDATFH 127
QY 414 TNVVRNNGSLYVPPGIFKSTCKIDITWFFDDQRCMKFGSWTYDGFQDLQLOQDETG 473
DB 128 TNVLNNSGHCQYLPFGIFKSCYDVRWFFDVQHCXKLFKGSWSYGSWLDLQMQE--- 184
QY 474 GDISSYVLNGEWELLGVPGKENEIYNNCCPEYIDITFAIIRRTLYFFNLIIPCVLI 533
DB 185 ADISGYIPNGEWDLVGIPGKRSERFECCKEYPDVTFTVMRRRTLYYGLNLIIPCVLI 244
QY 534 ASMLLGLTLPDSEKLSLGVTLISLTVFLNVAETMPATSDAVL----- 581
DB 245 SALALLVFLPADSGEKISLGITVLLSLTVFMLLVAEIMPATSDVPLIAQYFASTMIIV 304
QY 582 -----WIRIVLCWLPWILRMSRPG-----RPLILEFFPTTP 612
DB 305 GLSVVTVTVIVLYQHHDDPGCKMPKWTETVILLNWCWFLMKRPGEDKVRP-ACQHKQR 363
QY 613 CSDTSSERKHQILSDVELKERSKSLANVLIDDDFRH-----NCRPMTPGGTLPHNPAF 668
DB 364 CSLASVE-----MSAVAPPPASNGNLLY-----IGFGLDGVHCV-TP----- 401
QY 669 YRTVYGGDDGSGIGPIGSTRMPDAVTH-----HTCIKSTVEYELGLILKEIRFTDQLRK 723

This Page Blank (usp10)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 06:45:49 ; Search time 8474.79 Seconds
(without alignments)
3938.047 Million cell updates/sec

Title: US-09-303-232-2

Perfect score: 4043

Sequence: 1 MKNAQLKLTVDDELMLAV.....MFAIIATIAVLISAPHIIVS 770

Scoring table: BLOSUM62

Xgapop 10.0 , Xgapext 0.5

Ygapop 10.0 , Ygapext 0.5

Fgapop 6.0 , Fgapext 7.0

Delop 6.0 , Delext 7.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+ p2n.model -DEV=xln

-Q=/cgn2_1/USPTO.spool/US09303232/runat_07052004_101110_23883/app_query.fasta_1.2261

-DB=GenEmbl -QFMT=fastap -SUFFIX=rge -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0

-UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45

-DOALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL

-OUTFMT=ptc -NORM=ext -HEAPSIZ=500 -MINLEN=0 -MAXLEN=2000000000

-USER=US09303232@cgn_1_1_8225@runat_07052004_101110_23883 -NCPU=6 -ICPU=3

-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG

-DEV TIMEOUT=120 -WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6

-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4043	100.0	2886	6	E58346 Nucleic aci
2	4043	100.0	2886	6	AX009610 Sequence
3	4014.5	99.3	2907	3	AF272778 Drosophil
4	3804.5	94.1	2834	3	AY036613 Drosophil
5	1991	49.2	1683	3	DME554210
6	1850	45.8	3629	3	AF143846
7	1850	45.8	3700	6	AX009612
8	1850	45.8	3701	6	E58347
9	1625.5	40.2	2023	3	AF321446
10	1622.5	40.1	2023	3	AF321445
11	1621.5	40.1	1665	3	DME554209
12	1621.5	40.1	1699	3	BT011147
13	1620.5	40.1	2023	3	AF321447
14	1611	39.8	2110	3	AF321449
15	1609	39.8	3029	3	AF143847
16	1609	39.8	3109	6	E58348
17	1609	39.8	3109	6	AX009614
18	1602	39.6	2068	3	AF321448
19	1567.5	38.8	2116	3	AY036614
20	1456.5	36.0	42079	2	AC017620
21	1456.5	36.0	158758	3	AC092243
22	1456.5	36.0	186803	3	AC092225
23	1456.5	36.0	272521	3	AE003642
24	1456.5	36.0	320754	3	DROSADH02
25	1199	29.7	1719	5	AY247962
26	1185.5	29.3	2088	10	AF225980
27	1183.5	29.3	2106	10	RATNARAD
28	1182.5	29.2	1848	10	MUSNARS
29	1179.5	29.2	2037	5	GGNNARA7A
30	1179.5	29.2	2696	5	GGY7NAREC
31	1174	29.0	1542	5	AY299466
32	1173.5	29.0	3030	10	S53987
33	1170.5	29.0	2090	5	GGANAREC
34	1163	28.8	1668	5	AY298752
35	1160.5	28.7	1551	4	BTAT
36	1160.5	28.7	2087	9	HSARA7A
37	1159.5	28.7	1509	9	AX054567
38	1159.5	28.7	1509	9	HSNACHRA7
39	1159.5	28.7	1876	6	AR055255
40	1159.5	28.7	1876	6	AR071403
41	1159.5	28.7	1876	6	AR173187
42	1159.5	28.7	1876	6	AR224030
43	1159.5	28.7	1876	6	AR282833
44	1159.5	28.7	1876	6	AX719088
45	1159.5	28.7	1876	9	HSU62436

ALIGNMENTS

RESULT 1

E58346
LOCUS E58346 2886 bp DNA linear PAT 18-JUN-2001
DEFINITION Nucleic acid encoding insect acetyl choline receptor subunit.
ACCESSION E58346
VERSION E58346.1 GI:13019345
KEYWORDS JP 2000023680-A/1.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 2886)
REFERENCE Martin, A., Nadja, E. and Thomas, S.
AUTHORS Nucleic acid encoding insect acetyl choline receptor subunit
TITLE Patent: JP 2000023680-A 1 25-JAN-2000;
JOURNAL BAYER AG
COMMENT OS Drosophila melanogaster
PN JP 2000023680-A/1
PD 25-JAN-2000
PF 26-APR-1999 JP 1999118159
PR 04-MAY-1998 DE 19819829.9
PI MARTIN ADAMUTSUJESUKI, NADJA ERASU, THOMAS SCHULTE PC
C12N15/09, A01K67/033, C07K14/705, C07K16/28, C12N1/21, C12N5/10, PC
C12Q1/68,
PC G01N33/15, G01N33/50// (C12N1/21, C12R1:19), C12N5/00, C12N5/00 CC

EH Key Location/Qualifiers
FT CDS (372)..(2681).
1..2886
/organism="Drosophila melanogaster"
/mol_type="genomic DNA"
/db_xref="taxon:7227"

FEATURES
source

ORIGIN
Alignment Scores:
Pred. No.: 7,88e-286 Length: 2886
Score: 4043.00 Matches: 770
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-303-232-2 (1-770) x E58346 (1-2886)

Qy 1 MetLysAsnAlaGlnLeuLysLeuThrGluValAspAspGluLeuTrpLeuAlaVal 20
Db 372 ATGAAAATGCACAACTGAACTGACTGAACTTGACGATGATGAGCTGTGGCTGGCAGTA 431

Qy 21 ArgLeuAlaHisCysSerSerAsnPheSerSerSerSerSerSerSerSerSerSer 40
Db 432 AGATTAGCGCACTGAGCAGCACTTTAGCAGCAGTAGCAGCACAAAGAACCAACAGCAGC 491

Qy 41 AsnGlnArgHisAsnGlnGlnLeuThrThrLeuGlnProArgSerLeuSerThrLysHis 60
Db 492 AACGAGGCGCACACCACTCACAACTGCAACCAAGGAGCTTAAGTACAAAACAC 551

Qy 61 HisSerAsnIleAlaSerGluGlnHisAsnSerGlnGlnGlnGluProAlaSerLysAsp 80
Db 552 CACAGCAACATTGCAAGCGAGCAGCACAAATAGCCAGCAACAGGAGCCAGCATCGAAGGAC 611

Qy 81 GluAspValAlaAsnHisGlyArgSerAsnAspGlnGlnThrHisLeuGlnLeuAsp 100
Db 612 GAGGATGTAGCAACCACTGAGAACCAATGACCCAGCAGAGCGCATCTGCAACAGCTAGAC 671

Qy 101 SerSerAsnMetLeuSerProLysThrAlaAlaAlaAlaThrAlaAlaGlyAspGluAla 120
Db 672 AGCAGCAACATGTTGTCGCAAAAGACAGCCGAGCAGCAACTGCTGCGCGGCGATGAAGCA 731

Qy 121 ThrThrGlnGlnProThrAsnIleArgLeuCysAlaArgLysArgGlnArgLeuArgArg 140
Db 732 ACAACCCCAACCAACCAACCAATAAGACTGTGTGTCAGCAAGCAACAGATTGCGTCGC 791

Qy 141 ArgArgLysArgLysProAlaThrProAsnGluThrAspIleLysLysGlnGlnLeu 160
Db 792 CGACGAAAAAGAAAAACCGCAACCCCAACGAAACAGATATCAAGAAACACACAGCACTT 851

Qy 161 SerMetProProPheLysThrArgLysSerThrAspThrTyrSerThrProAlaAlaThr 180
Db 852 AGCATGCTCCCTTCAAAACCGGCAAAATCCACGGACACCTACAGCACACACAGCAACA 911

Qy 181 ThrSerCysProThrAlaThrTyrMetGlnCysArgAlaSerAspAsnGluPheSerIle 200
Db 912 ACCAGCTGTCGACAGCACCTCATGCAATGTGCGACGCAATAGTTCAGTATTCAGTATT 971

Qy 201 ProfileSerArgHisAspArgValSerThrAlaThrPheAlaTrpValLeuHisValLeu 220
Db 972 CCGATATCGAGACATGATAGATATCCAGGCCACATTCGCTGGGTGTGTCATGTGCTG 1031

Qy 221 GlnValLeuLeuValSerLeuGlnGlnTrpGlnLeuHisValGlnGlnArgSerValLeu 240
Db 1032 CAGGTGCTGCTGCTGCTGCTGCAACAGTGGCAACTTCACGTGCAACAGCATCGGTGCTA 1091

Qy 241 LeuPheArgArgIleAlaAlaSerThrIleAlaPheIleSerTyrLeuGlySerPheAla 260
Db 1092 CTGTTTCAGAAAGGATCGCAGCGAGCACCATCGCCTTCATTTCTATTATTAGCAGCTTTGCA 1151

Qy 261 AlaGlnLeuLysAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSer 280
Db 1152 GCGCACTGAAAAATAGCAGCAGCAGCAGTAGCAGCAGCAACAGCAACAACAGCAGC 1211

Qy 281 ThrGlnIleLeuAsnGlyLeuAsnLysHisSerTrpIlePheLeuLeuIleTyrLeuAsn 300
Db 1212 ACGCAAAATATTAAACGGACTTAATAAACACTCATGGATATTTTATTGATATATTGAAAT 1271

Qy 301 LeuSerAlaLysValCysLeuAlaGlyTyrHisGluLysArgLeuLeuHisAspLeuLeu 320
Db 1272 TTATCTGCTAAAGTTTGCCTAGCAGGATATCATGAAAGAGACTGTTACAGATCTTTTG 1331

Qy 321 AspProTyrAsnThrLeuGluArgProValLeuAsnGluSerAspProLeuGlnLeuSer 340
Db 1332 GATCCTTATAATACACTAGAACGTCCTGCTCAATGAATCGGACCCGTTACAAATTAAGC 1391

Qy 341 PheGlyLeuThrLeuMetGlnIleAspValAspGluLysAsnGlnLeuLeuValThr 360
Db 1392 TTTGGTTTAACTTTAATGCAAAATTTATCGATGCGACGAGAAAAATCAATTCTGACTACT 1451

Qy 361 AsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspTyr 380
Db 1452 AATGTGGTTTAAACCTGGAGTGAACAGACATGATCTCCGCTGGAAACCTCCGACAT 1511

Qy 381 GlyGlyValLysAspLeuArgIleProProHisArgIleTrpLysProAspValLeuMet 400
Db 1512 GCGGAGTTAAGGATCTGCGAATACCCGCGCATCGCATCTGGAAGCCGACGTGCTGATG 1571

Qy 401 TyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsnValValArgAsn 420
Db 1572 TACAACAGTGGCGGATGAGGAGTTTGACGGCACTTACCAGACGAGCTGTGGTGGCGAAC 1631

Qy 421 AsnGlySerCysLeuTyrValProProGlyIlePheLysSerThrCysLysIleAspIle 440
Db 1632 AACGGCTGTGCTATACGTTCCCGCGGGGATCTTCAAGTCGAGCTGCAAGATCGACATC 1691

Qy 441 ThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThrTyrAsp 460
Db 1692 ACGTGGTTCCCTTCGATGACCGCGGTGCGAGATGAAGTTCCGCGAGTTGGACCTACGAC 1751

Qy 461 GlyPheGlnLeuAspLeuGlnLeuGlnAspGluThrGlyGlyAspIleSerSerTyrVal 480
Db 1752 GGAATTCAGCTGGATTTACAATTAAGATGAATGCGGTGATATCAGCAGTTTACGTTG 1811

Qy 481 LeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLysArgAsnGluIleTyrTyrAsn 500
Db 1812 CTCACGCGGAGTGGGAACCTACTGGGTGTGCCCGCAACCGTAACGAGATCTATTACAC 1871

Qy 501 CysCysProGluProTyrIleAspIleThrPheAlaIleIleIleArgArgArgThrLeu 520

```
Db 1872 TGCTGCCGGACCTATATAGACATCACCTTCCCATCATCTCCGCGAGCAACTG 1931
Qy 521 TyrTyrPhePheAsnLeullelleProCysValLeulleAlaSerMetAlaLeuLeuGly 540
Db 1932 TACTATTCTTCAACCTGATCATACCTTGTTACTGCTGCTCCATGGCCTTCTCGGA 1991
Qy 541 PheThrLeuProProAspSerGlyGluLeuSerLeuGlyValThrLeuLeuSer 560
Db 1992 TTCACTCTGCCCGCAGATTCCGGTGAAATAATTCGCTGGGTGTATCCATCTTCTCTCG 2051
Qy 561 LeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValPro 580
Db 2052 CTGACCGTGTTCTCGAATAGTTGCGGAGACATGCGGCTACTTCGATGCGGTGCCA 2111
Qy 581 LeuTrpIleArgIleValPheLeuCysTrpLeuProTrpIleLeuArgMetSerArgPro 600
Db 2112 TTGTGGATAGCATCGTGTGTTTTGCTGGCTGCATGATATTCGAATGAGTCGCCCA 2171
Qy 601 GlyArgProLeulleLeuGluPheProThrThrProCysSerAspThrSerSerGluArg 620
Db 2172 GGACGACCGCTGATCTTAGAGTTCCCGACACGCGCTGTTGGGACATCTCCGAGCGG 2231
Qy 621 LysHisGlnIleLeuSerAspValGluLeuLysGluArgSerSerLysSerLeuLeuAla 640
Db 2232 AAGCACCAGATACTCTCCGAGTTGAGCTGAAGAGCGCTCGTGAATCGCTCTGCC 2291
Qy 641 AsnValLeuAspIleAspAspAspPheArgHisAsnCysArgProMetThrProGlyGly 660
Db 2292 AACGTACTAGACATCGATGATGACTTCGGGCACAAATTTGTCGCCCATGACGCCCGCGGA 2351
Qy 661 ThrLeuProHisAsnProAlaPheThrArgThrValTyrGlyGlnGlyAspAspGlySer 680
Db 2352 AACTGCGCACACACCCGGCTTTCTATCGACGGTTTTATGACAAAGCGACGATGGCAGC 2411
Qy 681 IleGlyProIleGlySerThrArgMetProAspAlaValThrHisHisThrCysIleLys 700
Db 2412 ATTGGGCCAATTGGCAGCACCCGAATGCCGATGCCGTACCACCATCATAGTCATCAAA 2471
Qy 701 SerSerThrGluTyrGluLeuGlyLeulleLeuLysGluIleArgPheIleThrAspGln 720
Db 2472 TCATCAACTGAATATGAATAGGTTTAACTTAAAGGAAATTCGCTTATAACTGATCAG 2531
Qy 721 LeuArgLysAspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetVal 740
Db 2532 CTACGTAAAGATGACGAGTGCATGACATTCGCAATGATTTGGAAATTTGCAGCTATGGTC 2591
Qy 741 ValAspArgLeuCysLeullellePheThrMetPheAlaIleLeuAlaThrIleAlaVal 760
Db 2592 GTTGACAGACTGTGCCTTATCATATTCCAAATGTTCCCAATATTAGCCCAATAGCTGTA 2651
Qy 761 LeuLeuSerAlaProHisIlelleValSer 770
Db 2652 CTACTATCGGCACCATATTATTGTCTCG 2681
```

```
RESULT 2
AX009610 LOCUS 2886 bp DNA linear PAT 06-SEP-2000
DEFINITION Sequence 1 from Patent EP0962528.
ACCESSION AX009610
VERSION AX009610.1 GI:9996842
KEYWORDS
SOURCE
ORGANISM Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1
REFERENCE Adamczewski,M.D., Schulte,T.D. and Oellers,N.D.
AUTHORS Nucleic acids encoding acetylcholin-receptor subunits from insects
TITLE Patent: EP 0962528-A 1 08-DEC-1999;
JOURNAL BAYER AG (DE)
FEATURES Location/Qualifiers
```

source

```
1. 2886
/organism="Drosophila melanogaster"
/mol_type="unassigned DNA"
/db_xref="taxon:7227"
372..2684
/notes="unnamed protein product"
/codon_start=1
/protein_id="CAC07499.1"
/db_xref="GI:9996843"
/db_xref="RENTREMBL:CAC07499"
/translation="MKNAQLKLEVDDELWLAURLAHCSNFSNFSSTRTTSSNOER
NQQLTLPESLSLTKHSNIASECHNSQQEPASQKDEVDVANGHSNQDQTHLQLODSS
NMSPKTAATAAAGAEATTQQTNTKLCARKQRLLRRRRKRPATNETDITKQQQL
SMSPFKTKSTDTYSTPAATSCPTATYMQCRASDNFEPIPIGRHDRVSTATFAWVLH
VLQVLSLQOQWLHVQORSVLLFRRIAASIAPISYLGSFAAQKLNSSSSSSSS
NNSSTQTLNGLNKHSMIFLIYLNLSAKVCLAGVHEKRLHLDLDPYNTLERPVLS
DPLQGLFGLTMOI IDVDEKNQLLVNWLKLEWDMNLRWNTSDYGVKDLRTIPHR
IKPDLVLMYNSADGPDGTQTNVVRNNGSLYVPGIFKSTCKIDITWFFEDQRC
EMKPGSWTDGFDLDLQDETSGDISSYVNGEWELLGVFGRNEIYINCCPEPIYD
ITFAIIRRTLYFFENLIIPCVLIASMAALLGFTLPDPSGEKLSGLVTILLSTFLVN
MVAETPATSDAVPLMIRIVFLCWLPMILRMSRPLIILEFTTTCGSDTSSBRKHQI
LSDVELKRSKSLANVLDDDDFRHNCRPMTPGGTLPHNPAFYRTVYGGDGGTG
PIGSTMPDAVTHHTCIKSTSEYELGLILKEIRFITDQLRKDDDECNDIANDWKFAMVY
VDRCLLIIFTFMFAILATIALLSAPHIIVS"
```

ORIGIN

Alignment Scores:
Pred. No.: 7,886-286 Length: 2886
Score: 4043.00 Matches: 770
Percent Similarity: 100.00% Conservative: 0
Best Local Similarity: 100.00% Mismatches: 0
Query Match: 100.00% Indels: 0
DB: 6 Gaps: 0

US-09-303-232-2 (1-770) x AX009610 (1-2886)

```
Qy 1 MetLysAsnAlaGlnLeuLysLeuThrGluValAspAspGluLeuTrpLeuAlaVal 20
Db 372 ATGAAATAATCAACAACCTGAACTGACTGAAGTTGACGATGATGAGCTGTGGTCAGTA 431
Qy 21 ArgLeuAlaHisCysSerSerAsnPheSerSerSerSerSerSerThrArgThrSerSer 40
Db 432 AGATTAGCGCACTGACGACCAACTTTAGCAGCAGTAGCAGCAGCAAGAACCCAGCAGC 491
Qy 41 AsnGlnArgHisAsnGlnGlnLeuThrThrLeuGlnProArgSerLeuSerThrLysHis 60
Db 492 AACGAGAGGCACACCAACGCAACTCACAACTCAACCAAGGAGCTTAAGTACAAACAC 551
Qy 61 HisSerAsnIleAlaSerGluGlnHisAsnSerGlnGlnGlnGlnProAlaSerLysAsp 80
Db 552 CACAGCAACATTGCAAGCGCAGCAGCACATAGCCAGCAACAGGAGCCAGCATCGAAGGAC 611
Qy 81 GluAspValAlaAlaAsnHisGlyArgSerAsnAspGlnGlnThrHisLeuGlnGlnLeuAsp 100
Db 612 GAGGATGTAGCAACCAACCGGTAGAACATGACACGAGCAGCATCTCGAACAGTAGAC 671
Qy 101 SerSerAsnMetLeuSerProLysThrAlaAlaAlaAlaAlaAlaGlyAspGluAla 120
Db 672 AGCAGCAACATGTTGTGCGCAAGACAGCGCAGCAGCAACTGCTGCCGCGCATGAAGCA 731
Qy 121 ThrThrGlnGlnProThrAsnIleArgLeuCysAlaArgLysArgGlnArgLeuArgArg 140
Db 732 ACAACCCAAACCAACCAACATAGACTGTGTGTCAGCAAGCAGCAACAGATTGCGTCGC 791
Qy 141 ArgArgLysArgLysProAlaThrProAsnGluThrAspIleLysLysGlnGlnGlnLeu 160
Db 792 CGACGAAAAGAAAACACAGCAACCCCAACAGAAACAGATATCAAGAAACAAACAGCAACT 851
Qy 161 SerMetProProPheLysThrArgLysSerThrAspThrTyrSerThrProAlaAlaThr 180
Db 852 AGCATGCCCTCCCTTCAAAACGGGCAATCCACGACACCTACAGCACACCCAGCAGCAACA 911
Qy 181 ThrSerCysProThrAlaThrTyrMetGlnCysArgAlaSerAspAsnGluPheSerIle 200
```


FEATURES	Location/Qualifiers	Score:	4014.50	Matches:	770
		Percent Similarity:	95.42%	Conservative:	0
source	/organism="Drosophila melanogaster"	Best Local Similarity:	95.42%	Mismatches:	0
	/mol_type="mRNA"	Query Match:	99.30%	Indels:	37
gene	/db_xref="taxon:7227"	DB:	3	Gaps:	1
	/chromosomes="2"	US-09-303-232-2 (1-770) x AF272778 (1-2907)			
CDS	/map="2L: 34E4-5"	QY	1	MetLysAsnAlaGlnLeuLysLeuThrGluValAspAspAspGluLeuTrpLeuAlaVal	20
	/dev_stage="embryo"	Db	289	ATGAAAAATGCACAACTGAAACTGACTGAAGTTGACGATGATGAGCTGTGGCTGCAGTA	348
variation	1..2907	QY	21	ArgLeuAlaHisCysSerSerAsnPheSerSerSerSerSerSerSerThrArgThrThrSerSer	40
	/gene="nAcRalpha-34E"	Db	349	AGATTAGCGCACTGCGAGCAGCAACTTTAGCAGCAGTACGACGACCAAGAACCAACAGCAGC	408
variation	/note="ion channel; neurotransmitter transmembrane receptor"	QY	41	AsnGlnArgHisAsnGlnGlnLeuThrThrLeuGlnProArgSerLeuSerThrLysHis	60
	/codon_start=1	Db	409	AACCAGAGGCACAAACCAAGCAACTCAACACTGCAACCAAGGAGCTTAAGTACAAAAACAC	468
variation	/product="nicotinic acetylcholine receptor Dalpha5 subunit"	QY	61	HisSerAsnIleAlaSerGluGlnHisAsnSerGlnGlnGlnProAlaSerLysAsp	80
	/protein_id="AAM13390.1"	Db	469	CACAGCAACATTTGCAAGCAGCAGCACAATACCCAGCAACAGAGCCAGCATCGAAGGAC	528
variation	/db_xref="GI:20152840"	QY	81	GluAspValAlaAsnHisGlyArgSerAsnAspGlnGlnThrHisLeuGlnGlnLeuAsp	100
	/translation="MKNAQLKLTVEVDDDELAVRLAHCSNFSSTSTSTSSNORH	Db	529	GAGGATGTAGCAACACCGCTAGAACAAATACCAGCAGCAGCGCATCTGCAACAGCTAGAC	588
variation	NQQLTLQPRSLSTKHHSNIASEQHNSSQQQEPASQEDVANHGKSNDDQTHLQQLDSS	QY	101	SerSerAsnMetLeuSerProLysThrAlaAlaAlaAlaThrAlaAlaGlyAspGluAla	120
	NMLSPKTAARATAGAEATQQPTNIRKLCARKQRRRRKRPATPNEDTIKKQQOL	Db	589	AGCAGCAACATGTTGTCGCAAGACAGCGCAGCAGCAACTGCTGCCGCGGATGAAGCA	648
variation	SMPEFKRSTDIYTPAATTSCTPTATMQCRASDNFSPISRHDRVSTATFAWLH	QY	121	ThrThrGlnGlnProThrAsnIleArgLeuCysAlaArgLysArgGlnArgLeuArgArg	140
	VLQVLVSQQQQLHVQQRVLLFRRIAATIAFISYLGSFAAQLKNSSSSSSSNS	Db	649	ACNACCCACACCAACAAACATAAGACTGTGTGCACCAAGCAGCAACAGATTGGCTGCG	708
variation	NNSSTQINGLNKHSWIFLLIYNLAKVCLAGYHEKRLHLLDDPYNTLERPVINES	QY	141	ArgArgLysArgLysProAlaThrProAsnGlnThrAspIleLysLysGlnGlnGlnLeu	160
	DPLQLSGFLTLMQI IDVEKNQLLVTNWLKLEWDMNLEWNTSDYGGVKDLIPHR	Db	709	CGACGAAAAAGAAAAACCCAGCAACCCCAACGAAACAGATATCAAGAAACACAGCACTT	768
variation	IMKPDVMTNSADEGPDGTQTNVVRNNGSCLYVPPGIFKSTCKIDITWFPDDQRC	QY	161	SerMetProProPheLysThrArgLysSerThrAspThrThrSerThrProAlaAlaThr	180
	EMKFGSWTDGFDLQLODETGDISSYVINGEWELLGVGRNEIYINCCEPEYID	Db	769	AGCATGCCCTCCCTTCAMAAACGCGCAAAATCCACGGACACCTACAGCACACAGCAACA	828
variation	ITFAILIRRTLYFNLLIPCVLIASMLLFTLPDSEKLSLGVTTILSLTVFIN	QY	181	ThrSerCysProThrAlaThrTyrMetGlnCysArgAlaSerAspAsnGlnPheSerIle	200
	MVAETPATSADVPLGLTYFCIMFVASSVSTILILNYHNRNADTHEMSEWIRIV	Db	829	ACCAGCTGTCCAGCAGCCACCTACATGCAATGTGAGCCAGCGCAATGAGTTCAGTATT	888
variation	LCWPLILMRSRPRLLIEFPPTPCSDTSSEKHOILSDVELKRSKSLLANVLDI	QY	201	ProIleSerArgHisAspArgValSerThrAlaThrPheAlaTrpValLeuHisValLeu	220
	DDDFRNCMPTEGGTLEPHNPAFVRTVYGGDDGSGIPGSTRMPDVAVHTTCKIST	Db	889	CCGATATCCAGACATGATAGAGTATCCACGGCCACATTCGGCTGGTGTGTCATGTCTG	948
variation	EDELGLLEIRFITDLQRKDECIANDKWFAMVVDRLCLITFMFAILLIIVL	QY	221	GlnValLeuLeuValSerLeuGlnGlnTrpGlnLeuHisValGlnGlnArgSerValLeu	240
	LSAPHILVS"	Db	949	CAGGTGCTGCTCGTGTGCTGCAACAGTGGCAACTTCACGTGCACACAGCGATCGGTGCTA	1008
variation	373	QY	241	LeuPheArgArgIleAlaAlaSerThrIleAlaPheIleSerTyrLeuGlySerPheAla	260
	/gene="nAcRalpha-34E"	Db	1009	CTGTTCCAGNAGGATCGCAGCAGCACCATCGCCCTTCATTTCTATTATTAGGCAGCTTGCA	1068
variation	/note="results in deletion of asparagine and serine; compared to B allele"	QY	261	AlaGlnLeuLysAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer	280
	/replace=""	Db	1069	GGCGAACTGAAAAATAGCAGCAGCAGCAGTAGTAGCAGCAGCAACACAGCAACAACAGCAGC	1128
variation	1119..1120	QY	281	ThrGlnIleLeuAsnGlyLeuAsnLysHisSerTrpIlePheLeuLeuIleTyrLeuAsn	300
	/gene="nAcRalpha-34E"	Db	1129	ACGCCAAATATTAAACGGACTTAATAAAACACTCATGGATATTTTTTATTGATATATTGAAT	1188
variation	/note="results in asparagine to serine substitution; compared to B allele"	QY	301	LeuSerAlaLysValCysLeuAlaGlyTyrHisGlnLysArgLeuLeuHisAspLeuLeu	320
	/replace="g"	Db	1189	TTATCTGCTAAAGTTTTCCTAGCAGGATATCATGAAAGAGACATGTTTACACGATCTTTTG	1248
ORIGIN	9.64e-284	QY	321	AspProTyrAsnThrLeuGluArgProValLeuAsnGluSerAspProLeuGlnLeuSer	340
	Length: 2907				

Db	1249	GATCCTTATTAATACACTAGAACGTCCTCCGTTCTCAATGAATCGGACCGCTTACAATTAAGC	1308
Qy	341	PheGlyLeuThrLeuMetGlnIleAaspValaspGluIysAsnGlnLeuLeuValThr	360
Db	1309	TTTGGTTAACTTTAATGCAAAATATCGATGTGACGAGAAAAATCAATTGCTAGTCACT	1368
Qy	361	AsnValTrpLeuIysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspTyr	380
Db	1369	AATGTGTGGTTAAACCTGGAGTGAACGACATGAATCTCCGCTGGAAACACCTCCGACTAT	1428
Qy	381	GlyGlyValIysAspLeuArgIleProProHisArgIleTrpLysProAspValLeuMet	400
Db	1429	GGCGAGCTTAAGGATCTCGCAATACCGCGCATCGCATCTGGAAGCGGACGCTCGTATG	1488
Qy	401	TyrAsnSerAlaaspGluGlyPheAaspGlyThrTyrGlnThrAsnValValValArgAsn	420
Db	1489	TACAACAGTGGGATGGGATTTTGACGGCACCCTACCAGAGCAAGCTGGTGGTGGGAAC	1548
Qy	421	AsnGlySerCysLeuTyrValProProGlyIlePheLysSerThrCysLysIleAspIle	440
Db	1549	AACGGCTCGTGTCTATACGTTCCCGCGGGGATCTTCAAGTCGACGTGCAAGATCGACATC	1608
Qy	441	ThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerThrTrpTyrAsp	460
Db	1609	ACGTGGTTCCTCCCTTCGATGACCGACGCGTGCAGATGAAGTTCCGCGAGTTGGACCTACGAC	1668
Qy	461	GlyPheGlnLeuAspLeuGlnLeuGlnAspGlnThrGlyGlyAspIleSerSerTyrVal	480
Db	1669	GGATTCGACGTGGATTTACAATTAACAAGATGAATCGCGGTGATATCAGCAGGTACGTG	1728
Qy	481	LeuAsnGlyGluTrpGluLeuGlyValProGlyLysArgAsnGluIleTyrTyrAsn	500
Db	1729	CTCAACGGCGAGTGGGAACCTACTGGGTGTGCGCGGCAACGTAACGAGATCTATTACAC	1788
Qy	501	CysCysProGluProTyrIleAaspIleThrPheAlaIleIleIleAArgArgThrIleu	520
Db	1789	TGCTGCCGGAACCTATATAGACATCACCTTCGCCATCATCTCCGCGGACGAACACTG	1848
Qy	521	TyrTyrPhePheAsnLeuIleIleProCysValleuIleAlaSerMetAlaLeuLeuGly	540
Db	1849	TACTATTCTTCAACCTGATCATACCTTGTGTACTGTGATTCCTCATGGCCCTTGTCTCGGA	1908
Qy	541	PheThrLeuProProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeuLeuSer	560
Db	1909	TTCACCTTCGCGCAGATTCGGGTGAAAAATATCGCTGGGTGTACCATCTTCTCTCTCG	1968
Qy	561	LeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValPro	580
Db	1969	CTGACCGTGTCTCTGAATATGTTGCCGAGACAATGCCGGCTACTTCCGATCGCGTGCCA	2028
Qy	581	Leu-----	581
Db	2029	TTGCTGGGTACATATTTCAATGTCATAATGTTATGTTAGTCTTCATCCGTTGTGTCAACG	2088
Qy	582	-----	582
Db	2089	ATTTTATATTAATATCATCATCGAAATGCTGATACGCACGAAATGTCGGAATGGATA	2148
Qy	584	ArgIleValPheLeuCysTrpLeuProTrpIleLeuArgMetSerArgProGlyArgPro	603
Db	2149	CGCATCGTGTCTTGTGCTGCTGCATGGATATTGCGAATGAGTCGCGCCAGGACGACCG	2208
Qy	604	LeuIleLeuGluPheProThrThrProCysSerAspThrSerSerGluArgLysHisGln	623
Db	2209	CTGATCTTAGATCTCCGACACACGCTTTCGGACACATCTTCGACGGGAAGACACAG	2268
Qy	624	IleLeuSerAspValGluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeu	643
Db	2269	ATACTCTCCGACGTTGAGCTGAAAGAGCGCTCGTCGAAATCGCTGCTGCCCAACGTA	2328
Qy	644	AspIleAspAspPheArgHisAsnCysArgProMetThrProGlyGlyThrLeuPro	663
Db	2329	GACATCGATGATGACTTCGGCGCAAAATTGTGCCCCCATGACGCCCGCGGAACACTGCCA	2388
Qy	664	HisAsnProAlaPheTyrArgThrValTyrGlyGlnGlyAspAspGlySerIleGlyPro	683
Db	2389	CAACCCCGCTTTCTATCGCACGGTTTATGGACAAGCGCAGATGGCAGCATTTGGGCCA	2448
Qy	684	IleGlySerThrArgMetProAspAlaValThrHisThrCysIleLysSerSerThr	703
Db	2449	ATTGGCAGCACCCGATGCGGATGCGGTCAACCATCATACGTGCATCAATCATCAACT	2508
Qy	704	GluTyrGluLeuGlyLeuIleLeuLysGluIleArgPheIleThrAspGlnLeuArgLys	723
Db	2509	GAATATGAATTAGCTTTAATCTTAAAGCAAAATTCGCTTATACTGATCAGCTACGTA	2568
Qy	724	AspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetValValAspArg	743
Db	2569	GATGACGAGTGCATGACATTCGCAATGATTTGMAATTTGCAGTATGTTGTTGACAGA	2628
Qy	744	LeuCysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSer	763
Db	2629	CTGTGCTTATCATATTCAATGTTCCGAATATTAGCCACAATAGCTGTACTACTATCA	2688
Qy	764	AlaProHisIleIleValSer 770	
Db	2689	GCACCATATATTGTTCTCG 2709	
LOCUS	AY036613	2834 bp mRNA linear INV 30-APR-2002	
DEFINITION	Drosophila melanogaster nicotinic acetylcholine receptor Dalphas subunit (nAcRalpha34E) mRNA, nAcRalpha34E-B allele, complete cds, alternatively spliced.		
ACCESSION	AY036613		
VERSION	AY036613.1 GI:20340268		
KEYWORDS	Drosophila melanogaster (fruit fly)		
SOURCE	Drosophila melanogaster		
ORGANISM	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.		
REFERENCE	1 (bases 1 to 2834) Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B. Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalphas5, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing		
AUTHORS	Genetics 160 (4), 1519-1533 (2002)		
TITLE	21969411 11973307		
JOURNAL	2 (bases 1 to 2834)		
MEDLINE	Grauso, M. and Sattelle, D.B.		
PUBMED	Direct Submission		
REFERENCE	Submitted (26-MAY-2001) MRC-FGU Human Anatomy and Genetics, University of Oxford, South Parks Road, Oxford OX1 3QX, UK		
AUTHORS	Location/Qualifiers		
TITLE	1. .2834		
JOURNAL	/organism="Drosophila melanogaster"		
MEDLINE	/mol_type="mRNA"		
PUBMED	/db_xref="taxon:7227"		
REFERENCE	/chromosome="2"		
AUTHORS	/map="34E4-34E5"		
TITLE	/dev_stage="embryo"		
JOURNAL	1. .2834		
FEATURES	/gene="nAcRalpha34E"		
source	/allele="B"		
	289 .1737		
gene	/gene="nAcRalpha34E"		
CDS	/note="ion channel; neurotransmitter transmembrane receptor; exon 5 is excluded due to exon skipping, which generates a loss of reading frame and a truncated polypeptide; alternatively spliced"		
	/codon_start=1		
	/product="nicotinic acetylcholine receptor Dalphas5"		

subunit"
 /protein_id="AAK67256.1"
 /db_xref="GI:20340269"
 /translation="MKNAQLKLTVEVDDDELMLAVRLAHCNSNISSSSSTRRTSSNQHRH
 NQUTITLQPRSLSTKHSNIASQEQHNSQQEPASKDEVDVANHGRSNDQOQTHLQOLDSS
 NMLSPKTAATAAAGDEATTOOPTNIRLCARQRRLRRRRKRPATNETIDIKKQOOL
 SMPFKTKSTDTYSTPAATISCTPATYMCRCASDNFSPISRHRDVSATATFAWLH
 VLOVLLSQOQLHVOORSVLLFRFAASTIAFISYLGSPAOLRSSSSSSSSNS
 GNSSTQLNGKHSMTFLILYNLSAKVCLAGVHEKRLHLHLLDPYNTILRPLVNE
 SDPLQLSGLTLMOIIVDVEKNQLLVYNWKLKWNMDNLRNITSDTGGVADLRIPPH
 RIWFDVLMYNSADEGFGDTYQTNVVRNNGSLYVPPGIFPKSTCKIDITWFFPDDOR
 CEMRFGSMTYDGFQVCPANVTRSIITAARNP"

ORIGIN

Alignment Scores:

Pred. No.: 2,09e-268 Length: 2834
 Score: 3804.50 Matches: 738
 Percent Similarity: 91.71% Conservative: 3
 Best Local Similarity: 91.34% Mismatches: 4
 Query Match: 94.10% Indels: 64
 DB: 3 Gaps: 3

US-09-303-232-2 (1-770) x AY036613 (1-2834)

Qy 1 MetLysAsnAlaGlnLeuLysLeuThrGluValAspAspGluLeuTrpLeuAlaVal 20
 Db 289 ATGAAAAATGCACAACTGAACTGACTGAACTTCACGATGATGAGCTGTGGCTGGCAGTA 348
 Qy 21 ArgLeuAlaHisCysSerSerAsnPheserSerSerSerThrArgThrThrSerSer 40
 Db 349 AGATTAGCGCACTCCAGCAGCAACATCAGCAGCAGTAGTAGCAGCACAAAGAACCCAGCAGC 408
 Qy 41 AsnGlnArgHisAsnGlnGlnLeuThrThrLeuGlnProArgSerLeuSerThrLysHis 60
 Db 409 AACAGAGGCACACACAGCAACTCACCAACTGCAACCAAGGAGCTTAAGTACAAAAAC 468
 Qy 61 HisSerAsnLeuAlaSerGluGlnHisAsnSerGlnGlnGlnGluProAlaSerLysAsp 80
 Db 469 CACAGCAACATTCAGAGCAGCAGCACAAATAGCCAGCAACAGAGCCAGCATCGAAGGAC 528
 Qy 81 GluAspValAlaAsnHisGlyArgSerAsnAspGlnGlnThrHisLeuGlnGlnLeuAsp 100
 Db 529 GAGGATGTAGCCAAACCAACGAGTAGAAGCAATGACAGCAGCAGCGCATCTGCAACAGCTAGAC 588
 Qy 101 SerSerAsnMetLeuSerProLysThrAlaAlaAlaAlaThrAlaAlaGlyAspGluAla 120
 Db 589 AGCAGCAACATGTTGTCCCAAGACAGCCGAGCAACTGTGCTCCGCGCATGAAGCA 648
 Qy 121 ThrThrGlnGlnProThrAsnIleArgLeuCysAlaArgLysArgGlnArgLeuArgArg 140
 Db 649 ACAACCCCAACCAACAAACATAAGACTGTGTGTCAGCGCAGCAGCAACGATTGCGTGC 708
 Qy 141 ArgArgLysArgLysProAlaThrProAsnGluThrAspIleLysLysGlnGlnGlnLeu 160
 Db 709 CGACGAAAGAAAGAACCCAGCAACCCCAACAGCAACAGATATCAAGAAACCAACAGCAACT 768
 Qy 161 SerMetProProPheLysThrArgLysSerThrAspThrThrThrSerProAlaAlaThr 180
 Db 769 AGCATGCTCCCTTCAAAAGCGCAATTCACGACACCTTACAGCAGCAGCAGCAGCAATA 828
 Qy 181 ThrSerCysProThrAlaThrTyrMetGlnCysArgAlaSerAspAsnGluPheSerIle 200
 Db 829 ACCAGCTGTCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 888
 Qy 201 ProIleSerArgHisAspArgValSerThrAlaThrPheAlaTrpValLeuHisValLeu 220
 Db 889 CCAGATATCGAGACATATAGATATCCAGCCACCATTCGCTGGGTGTGATGTGCTG 948
 Qy 221 GlnValLeuValSerLeuGlnGlnTrpGlnLeuHisValGlnGlnArgSerValLeu 240
 Db 949 CAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1008
 Qy 241 LeuPheArgArgIleAlaAlaSerThrIleAlaPheIleSerTyrLeuGlySerPheAla 260

Db 1009 CTGTTTCAAGAGATCGCAGCAGCACCATCCCTTCAATTCCTATTTAGGCAGCTTTGCA 1068
 Qy 261 AlaGlnLeuLysAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 279
 Db 1069 GCGCAACTGAGGAGCAGCAGCAGTAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 1128
 Qy 280 SerThrGlnIleLeuAsnGlyLeuAsnLysHisSerThrIlePheLeuLeuIleTyrLeu 299
 Db 1129 AGCAGCAAAATATTAACCGCATTAATAAACACCTCATGATATTTTATTGATATATTG 1188
 Qy 300 AsnLeuSerAlaLysValCysLeuAlaGlyTyrHisGluLysArgLeuLeuHisAspLeu 319
 Db 1189 AATTATCTGCTAAAGTTTGCCTAGCAGGATATCATGAAAAGAGACTGTTACAGCATTT 1248
 Qy 320 LeuAspProTyrAsnThrLeuGluArgProValLeuAsnGluSerAspProLeuGlnLeu 339
 Db 1249 TTGGATCTTATATACACAGCAGTCCCGTTCTCAATGATCGGACCCGTTACAATTA 1308
 Qy 340 SerPheGlyLeuThrLeuMetGlnIleAspValAspGluLysAsnGlnLeuLeuVal 359
 Db 1309 AGCTTTGGTTTAACTTTAATGCAATATATCATGTGAGCAGAAAAATCAATTGCTAGTC 1368
 Qy 360 ThrAsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAsp 379
 Db 1369 ACTAATGTGTGTTAAACCTGGAGTGAACGACATGAATCTCCGCTGGAACACCTCCGAC 1428
 Qy 380 TyrGlyGlyValLysAspLeuArgIleProProHisArgIleTyrLysProAspValLeu 399
 Db 1429 TATGGCGAGTTAAGGATCTGCCAATACCGCCGATCGCATCTGGAAGCCGAGCTGCTG 1488
 Qy 400 MetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsnValValArg 419
 Db 1489 ATGTACAAACAGTCGGATGAGGATTTGACGGCACCTTACCAGACGAACGTGGTGGTGG 1548
 Qy 420 AsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSerThrCysLysIleAsp 439
 Db 1549 AACACCGCTCGTGTCTATACGTTCCGCGGGGATCTTCAAGTCGACGTCGCAAGATCGAC 1608
 Qy 440 IleThrTrpPheProPheAspAspGlnArgCysGluMetLysPheGlySerThrTyr 459
 Db 1609 ATCAGCTGGTTCCTTCGATGACCGAGGTGGAGATGAAGTTCGCGAGTTGACCTAC 1668
 Qy 460 AspGlyPheGlnLeuAspLeuGlnLeuAspGluThrGlyGlyAspIleSerSerTyr 479
 Db 1669 GACGGATTCCA----- 1679
 Qy 480 ValLeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLysArgAsnGluIleTyrTyr 499
 Db 1680 -----GGTGTGCCCGGCAACGTAAACGAGATCTATTAC 1712
 Qy 500 AsnCysCysProGluProTyrIleAspIleThrPheAlaIleIleIleArgArgArgThr 519
 Db 1713 AACTGCTGCCGGAACCTATATAGACATCACTTCGCCATCATCATCCGCCGACGAAACA 1772
 Qy 520 LeuTyrTyrPhePheAsnLeuIleIleProCysValLeuIleAlaSerMetAlaLeuLeu 539
 Db 1773 CTGTACTATTTCTTCAACCTGATACCTTGTGTACTGATTCCTCCATCGCTTGGCTGCTC 1832
 Qy 540 GlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeuLeu 559
 Db 1833 GGAATCACCTGCGCGCAGATTGCGGTGAAAAATATATCGCTGGGTGTTTACCATCTTGCTC 1892
 Qy 560 SerLeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaVal 579
 Db 1893 TCGCTGACCGTGTCTCTGAAATATGTTGCGGAGACAAATCCGCTACTTCCGATGCGGTG 1952
 Qy 580 ProLeu----- 581
 Db 1953 CCATTGCTGGGTACATATTTCAATTGCAATAATGTTTATGGTAGTTCATCCGTTGTTGTC 2012
 Qy 582 -----Tyr 582


```
Db      2013  ACGATTTAATTAATTAATCATCATCGAAATGCTGATACGACGAAATGTCGCAATGG 2072
Qy      583  IleArgIleValPheLeuCysTyrLeuProTyrIleLeuArgMetSerArgProGlyArg 602
Db      2073  ATACGCATCGTGTGTTTGTGCTGCCATGGAATATGCGAATGAGTCGCCACGACGA 2132
Qy      603  ProLeuIleLeuGluPheProThrThrProCysSerAspThrSerGluArgLysHis 622
Db      2133  CCGCTGATCCTAGAGTTCGACACGACCCCTGTCGACACATCTCGACGCGAAGCAC 2192
Qy      623  GlnIleLeuSerAspValGluLeuLysGluArgSerSerLysSerLeuLeuAlaAsnVal 642
Db      2193  CAGATACTCTCCGAGGTTGAGCTGAAAGAGCGCTCGTCGAAATCGCTGCGCCAACTGA 2252
Qy      643  LeuAspIleAspAspPheArgHisAsnCysArgProMetThrProGlyGlyThrLeu 662
Db      2253  CTAGACATCGATGATGACTTCGCGCACAAATGTCGCCCATGACGCCCGCGGACACATG 2312
Qy      663  ProHisAsnProAlaPheTyrArgThrValTyrGlyGlnGlyAspAspGlySerIleGly 682
Db      2313  CCACACAAACCGCGTTCATCGACCGGTTTATGACACGCGTTATGACACGCGATGCGACGATGGG 2372
Qy      683  ProIleGlySerThrArgMetProAspAlaValThrHisHisThrCysIleLysSerSer 702
Db      2373  CCAATTGGCAGCACCCGAAATGCGGATGCGGTCCATCATCATCATCATCAATCATCA 2432
Qy      703  ThrGluTyrGluLeuGlyLeuIleLeuLysGluIleArgPheIleThrAspGlnLeuArg 722
Db      2433  ACTGAATATGAATAGGTTTAAATCTTAAGAGAAATTCGCTTTAATCACTGATCAGCTACGT 2492
Qy      723  LysAspAspGluCysAsnAspIleAlaAsnAspTyrLysPheAlaAlaMetValValAsp 742
Db      2493  AAAAGATGACGATGCAATGATGCGCAATGATGCGCAATGATGCGCAATGATGCGTGTAC 2552
Qy      743  ArgLeuCysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeu 762
Db      2553  AGACTGTGCTTATCATATTACAAATGTTGCAATATGTCGAATATGACCAATAGCTGTACTACTA 2612
Qy      763  SerAlaProHisIleIleValSer 770
Db      2613  TCAGACCACATATTATTGTCTCG 2636

RESULT 5
DME554210 1683 bp mRNA linear INV 01-APR-2003
LOCUS      Drosophila melanogaster mRNA for nicotinic acetylcholine receptor
DEFINITION subunit Dalp7 (nAcRalpha-18C gene).
ACCESSION  AJ554210
VERSION     AJ554210.1 GI:29466436
KEYWORDS    nAcRalpha-18C gene; nicotinic acetylcholine receptor subunit
SOURCE      Dalp7.
ORGANISM    Drosophila melanogaster (fruit fly)
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE   1 (bases 1 to 1683)
AUTHORS     Millar,N.S.
TITLE       Direct Submision
JOURNAL     Submitted (31-MAR-2003) Millar N.S., Department of Pharmacology,
            University College London, Gower Street, London, WC1E 6BT, UNITED
            KINGDOM
FEATURES    Location/Qualifiers
            1..1683
             /organism="Drosophila melanogaster"
             /mol_type="mRNA"
             /db_xref="taxon:7227"
            1..1683
             /gene="nAcRalpha-18C"
            26..1654
             /gene="nAcRalpha-18C"
             /codon_start=1
             /product="nicotinic acetylcholine receptor subunit
```

```
Dalp7a7"
/protein_id="CAD86936.1"
/db_xref="GI:29466437"
/translation="MSFPQPSLPEATANGRMIVYGLGLIMIPACAAGPHEKRLIH
ALLDNNLSLRPVVNSDPLQLSGLTLMQIIDVDEKNOLLITNWLKLEWDMNLSW
NSSEFGRLRIPPHRLWKPDMYNSADEGFDGTYATNVVVRNNGSCLYVPPGIFK
STKIDITWFDFDQRCMKFGSYDGLDQLQDEAGDISSTITNGEWDLGVP
GKENEIYNCCPEPIDITPAILIRKTLIYFFNLIIVPCVLIASMLLGLTLLPDSGE
KLSGLTLLSLTFLNMAETWPAATSDAYPLLGKFCNPFVSSSSASGEKKQIQN
HRNPDPHEKSEMRVIFLWLPCLRMORPGQVCEPCPPPPSSSSASGEKKQIQN
VEUKSSKSLLANVLDDDDFCNRHCASATLPHQTYTYRMYRQGDGSGVGPQA
GPVVDGGRHEAHSHTCLTSSAEYELAILKELRWITEQLKREDBTSDITRDNKPAAV
VDRLCGLIITFTIATLAVLSAPHFIVSGVRG"
26..131
/gene="nAcRalpha-18C"
132..1651
/gene="nAcRalpha-18C"
/product="nicotinic acetylcholine receptor subunit
Dalp7a7"

Alignment Scores:
Pred. No.:      3.85e-136      Length:      1683
Score:          1991.00        Matches:      382
Percent Similarity: 80.95%      Conservative: 43
Best Local Similarity: 72.76%    Mismatches:   46
Query Match:     49.25%        Indels:       54
DB:              3            Gaps:         8

US-09-303-232-2 (1-770) x DME554210 (1-1683)
Qy      295  LeuLeuIleTyr-----LeuAsnLeuSerAlaLysValCysLeuAlaGlyTyrHis 311
Db      80  ATGCTGGTCTATGCGCTGGGACTTTTAATTATGATACCGGCTTGTGCGCTGGACCCCAT 139
Qy      312  GluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeu 331
Db      140  GAGAAGCGGCTTACTCCAGCCCTTCTGGACAACTACACAGCCGTGGAGCGTCCGGTGGTC 199
Qy      332  AsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAspVal 351
Db      200  AATGAATCCGATCCATTGCCAATCGAGCTTCGAGCTACACCTCATCATCGATATTCATGTG 259
Qy      352  AspGluLysAsnGlnLeuLeuValThrAsnValThrLysLeuLysLeuGluTrpAsnAspMet 371
Db      260  GACGAAAGAAATCAACTGCTTATAACGAATATTTGGCTCAAAATGGAATGGAACGATATG 319
Qy      372  AsnLeuArgTrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIleProProHis 391
Db      320  AATCTTCGATGGAATTCGAGTGAGTTCGGTGTGTCGGGATCTCGAATTCGCCACAT 379
Qy      392  ArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThr 411
Db      380  CGCTATGAAACCGAGATGCTATGATGTAACAGCTGCGAGGAGGCTTCGATGGAACG 439
Qy      412  TyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProProGlyIle 431
Db      440  TAGCCCAAAATGTGGTTCGCAATTAATGGAGCTGTCTGTACGTACCGACGATATA 499
Qy      432  PheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGlu 451
Db      500  TTTAAGTCAACGTGTAAGATGACATTCGTTTCAATCGGATTCGAGATTCGAGATGTGAA 559
Qy      452  MetLysPheGlySerTyrTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAspGlu 471
Db      560  ATGAAATTTGGTTCGTGGACCTACGATGGGTTTCAGTTGGACCTGACGTTCGAGACGAA 619
Qy      472  ThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyValPro 491
Db      620  GCTGTGGCGACATTTCTAGCTTTATAACCAATGCGAATGGGACTTGTAGTGTGGCC 679
Qy      492  GlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPhe 511
```



```
QY 375 TrpAsnThrSerAspTyrGlyValLysAspLeuArgLleProProHisArgLleTrp 394
Db 623 TGGAACTTTCAGATTTCGGGGGGTCAAGATTTAAGAGTGCCACCCACAGACTATGG 682
QY 395 LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThr 414
Db 683 AAACAGACGCTCTTATGTACACAGCGCGAGCAAGAGGTTTCGACAGCACGTATCCAACG 742
QY 415 AsnValValValArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSer 434
Db 743 AACGTGGTGGTGGCGAACAACGGCTCGTGTCTACGTGCGCGCGCATCTCAAGAGC 802
QY 435 ThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPhe 454
Db 803 ACCTGCAAGATCGACATCACTCGTTCCCTTCACAGCACCAACGATGCGAGATGAAGTTT 862
QY 455 GlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuGlnAspGluThrGlyGly 474
Db 863 GGCAGCTGGACTTATGATGTTTATCAGTTGGATCTACAACCTACAGGATGAAGGGCGGA 922
QY 475 AspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuGlyValProGlyLysArg 494
Db 923 GATATAAGCGATTITGTACGAATGGCGAATGGAGTTAATAGGAGTCCCGCGCAAGCGC 982
QY 495 AsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaIleIle 514
Db 983 AACGAGATCTACTACAACTGTTGTCGGAGGCACTACATCGACATCAGTTTTCGGTGGTG 1042
QY 515 IleArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysValLeuIleAla 534
Db 1043 ATCGGAGGAAACGCTCTACTTCTTCAATCTGATCGGCCCTCGGTGCTCATCGCC 1102
QY 535 SerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGly 554
Db 1103 TCCAAGCTCTATTGGGGTTTCACTTCCCTCCGACTCCGGAGAAAAGTTGCTCTTAGGT 1162
QY 555 ValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAla 574
Db 1163 GTGACGATATTACTGTGTTGAGCGGTCTTCTACATGTTGGCGGAGAGATGCCAGCG 1222
QY 575 ThrSerAspAlaValProLeu----- 581
Db 1223 ACCTCGAGCGCGTGCCTTGTCTCGGCACCTTCTCACTGCATCAGTTTCATGGTGGCT 1282
QY 581 ----- 581
Db 1283 TCCTCGCTGCTCCACCACTGATCCTCAACTACCAACCGCGCAGCAGACACTCAC 1342
QY 582 -----TrpIleArgIleValPheLeuCysTrpLeuProTrpIleLeuArgMet 597
Db 1343 GAAATGAGTGATTGGATTCTGTCGTGTTCTTTATTGGCTGCGGGTGGCGGATG 1402
QY 598 SerArgProGly-----ArgProLeuIleLeuGluPheProThrThrProCysSer 614
Db 1403 TCACGCGCGGCTCGGCGAGCAGCGCGCGCGGTACCTTCGCGCGCG----- 1456
QY 615 AspThrSerSerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGluArgSer 634
Db 1457 -----GACTGGAGCTGCGCGAGCGCTCC 1480
QY 635 SerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsnCysArg 654
Db 1481 TCCAAGTCGCTCTAGGAAGCTGCTCGACATCGATCGACTTCGCCAC----- 1531
QY 655 ProMetThrProGlyThrLeuProHisAsnProAlaPheTyrArgThrValTyrGly 674
Db 1532 -----CCGCAAGCGCAGCAGCGCAATGTCGCGATCTACAG----- 1570
QY 675 GlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAspAlaValThr 694
Db 1571 GGGGGTGAGGAGAAATGCGCGGG-----TTGGCG 1600
```

```
QY 695 HisHisThrCysIleLysSerSerThrGluTyrGluLeuGlyLeuIleLysGluIle 714
Db 1601 CGGCACAGATTCTTTC-----GGTGTGCAGCTACAGCTCTCCCTCATTTCTGAAGGAGATT 1654
QY 715 ArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsnAspIleAlaAsnAspTrp 734
Db 1655 AGAGTCATCATCAGATCATAGATCGCAAGACACCAAGATGCGGACATTTCGCGCACTGG 1714
QY 735 LysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrMetPheAlaIle 754
Db 1715 AAGTTTCGCGGCATGTCGTGGACAGACTGTCCTTATTATCTTTACCTCTTCAATC 1774
QY 755 LeuAlaThrIleAlaValLeuLeuSerAlaProHisIleIleValSer 770
Db 1775 ATCCCAAGCTAGCTGCTGCTGTCGCGCCACACATCATGCTGTCG 1822
RESULT 8
LOCUS E58347 3701 bp DNA linear PAT 18-JUN-2001
DEFINITION Nucleic acid encoding insect actyl choline receptor subunit.
ACCESSION E58347
VERSION E58347.1 GI:13019346
KEYWORDS JP 2000023680-A/2.
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE
1 (bases 1 to 3701)
AUTHORS Martin, A., Nadja, E. and Thomas, S.
TITLE Nucleic acid encoding insect actyl choline receptor subunit
JOURNAL Patent: JP 2000023680-A 2 25-JAN-2000;
BAYER AG
COMMENT
OS Heliothis virescens
PN JP 2000023680-A/2
PD 25-JAN-2000
PF 26-APR-1999 JP 1999118159
PR 04-MAY-1998 DE 19819829.9
PI MARTIN ADAMUTSUEUSUKI,NADJA ERASU,THOMAS SCHULTE PC
C12N15/09,A01K67/033,C07K14/705,C07K16/28,C12N1/21,C12N5/10, PC
C12Q1/68.
PC G01N33/15,G01N33/50//(C12N1/21,C12R1/19),C12N15/00,C12N5/00 CC
FEATURES
FH Key Location/Qualifiers
FT CDS (335)..(1822).
source
1..3701
/organism="Heliothis virescens"
/mol_type="genomic DNA"
/db_xref="taxon:7102"
ORIGIN
Alignment Scores:
Pred. No.: 2,21e-125 Length: 3701
Score: 1850.00 Matches: 361
Percent Similarity: 75.78% Conservative: 30
Best Local Similarity: 69.96% Mismatches: 49
Query Match: 76
DB: 6 Gaps: 7
US-09-303-232-2 (1-770) x E58347 (1-3701)
QY 295 LeuLeuIleTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyrHisGluLysArg 314
Db 383 CTGCTGCTGCTGCTGCTGCTGCTGCGCGAGGGGGGACGCTGCGGGTACCACGAGACGG 442
QY 315 LeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeuAsnGluSer 334
Db 443 CTACTGCACCACTATTGGACCACCTACAACGCTACTGGAGAGCCCGCTCGTCAACGAGC 502
QY 335 AspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleIleAspValAspGluLys 354
Db 503 GACCCGCTGACGCTCTCCCTTCGGCTCAGCTCATCGATCATCGCTGACGAGAG 562
```

Qy	355	AsnGlnLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArg	374
Db	563	AACAGCTTTTAAATAACAAACATCTGGCTAAACATAGAGTGAATGATATGAACCTTAGG	622
Qy	375	TrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIleProProHisArgIleTrp	394
Db	623	TGGAACACTTCAGATTTCGGCGGGTCAAGATTTAGAGTGGCCACCCACAGACTATGG	682
Qy	395	LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThr	414
Db	683	AAACAGACGCTCCTTATGTACACAGCGCGGACGAGGGTTCGACAGCATATCCAAACG	742
Qy	415	AsnValValAlaArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSer	434
Db	743	AACGTGGTGGTGGGAAACACCGCTCGTGTCTGTACGTGGCGCCCGGCACTCTCAAGAC	802
Qy	435	ThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPhe	454
Db	803	ACCTGCAAGATCGACATCACCTGGTTCCCTTCGACGACCAACAGATCGAGATGAAGTTT	862
Qy	455	GlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuGlnAspGluThrGlyGly	474
Db	863	GGCAGCTGGACTTATGATGGTTATCAGTTGGATCTACAACCTACAGGATGAAGCGGGCGGA	922
Qy	475	AspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLysArg	494
Db	923	GATATAGCAGCTTTGTTCAGAAATGGCAATGGGAGTTAATAGGATCCCGCGCAGCGC	982
Qy	495	AsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaIleIle	514
Db	983	AACAGATCTACTACAACCTGTTGTCGGAGCCATACATCGACATCACGTTTGGCGTGTG	1042
Qy	515	IleArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysValLeuIleAla	534
Db	1043	ATCCGAGGAAACGCTACTACTTCTTCAATCTGATCGTGCCTCGCTCATCGCC	1102
Qy	535	SerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGly	554
Db	1103	TCCATGGCTCTATTGGGGTTTCCCTTGCCTCCAGACTCCGGAGAAAGTTGCTTTAGGT	1162
Qy	555	ValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAla	574
Db	1163	GTGAGGATATTACTGCTGTGACGGTGTCTCAACATGGTGGGAGACGATGCCAGCG	1222
Qy	575	ThrSerAspAlaValProLeu	581
Db	1223	ACGTGGGACGGCGTGCCTTGTCTGGCACCTACTTCAACTGCATCATGTTTCATGGTGCT	1282
Qy	581	-----	581
Db	1283	TCCTCCGTGCTCCACCATACTGATCTCAACTACCAACCCGCGCAGCAGACTCAC	1342
Qy	582	-----	582
Db	1343	GAATGAGTGAATGGATTGTTGGTGTGTTTATTTGCTGCTGGTGGTGGCGCATG	1402
Qy	598	SerArgProGly	614
Db	1403	TCACGGCCCGCTCGCGACGACGCGCGCGCGCGCGGTACCTCCGCGCGCG	1456
Qy	615	AspThrSerSerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGluArgSer	634
Db	1457	-----	1457
Qy	635	SerLysSerLeuLeuAlaAsnValLeuAspIleAspAspAspPheArgHisAsnCysArg	654
Db	1481	TCCAAAGTCGCTCCTAGCAACGCTGCTCGACATCGATGACGACTTCGCGCAC	1531
Qy	655	PrometThrProGlyGlyThrLeuProHisAsnProAlaPheTyrArgThrValTyrGly	674
Db	1532	-----	1570

Qy	675	GlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAspAlaValThr	694
Db	1571	GGGGGTGAGGAGAAATGGCGCGGG-----	1600
Qy	695	HisHisThrCysIleLysSerSerThrGluTyrGluLeuGlyLeuIleLysGluIle	714
Db	1601	GCACACAGTTGCTTC-----GGTGTGACTACGAGCTCTCCCTCATCTTGAAGAGAT	1654
Qy	715	ArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsnAspIleAlaAsnAspTrp	734
Db	1655	AGAGTCATCATCATGATCATGCGCAGGACGACGAGATCGGACATTCGCGGACATG	1714
Qy	735	LysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrMetPheAlaIle	754
Db	1715	AAGTTGCGCGCCATGGTGGGACAGACTGTGCTTATTATCTTACCTGTTTCAACATC	1774
Qy	755	LeuAlaThrIleAlaValLeuLeuSerAlaProHisIleIleValSer	770
Db	1775	ATCGCCAGCTAGCCGCTGCTGTCGCGCCACACATCATGTTGTCG	1822

RESULT	9
LOCUS	AF321446
DEFINITION	Drosophila melanogaster nicotinic acetylcholine receptor Dalpha6 subunit variant type II (nAChRalpha-30D) mRNA, complete cds, alternatively spliced.
ACCESSION	AF321446
VERSION	AF321446.1
KEYWORDS	GI:20152846
SOURCE	Drosophila melanogaster (fruit fly)
ORGANISM	Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE	1. (bases 1 to 2023) Grauso, M., Keenan, R.A., Culetto, E. and Sattelle, D.B. Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing Genetics 160 (4), 1519-1533 (2002)
JOURNAL	21969411
MEDLINE	11973307
PUBMED	2 (bases 1 to 2023)
REFERENCE	Grauso, M. and Sattelle, D.B. Direct Submission Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional Genetics Unit, University of Oxford, South Park Road, Oxford OX1 3QX, UK
AUTHORS	
TITLE	
JOURNAL	
FEATURES	Location/Qualifiers 1. .2023 /organism="Drosophila melanogaster" /mol_type="mRNA" /db_xref="taxon:7227" /chromosome="2" /map="30D1" /dev_stage="embryo" 1. .2023 /gene="nAChRalpha-30D" 379. .1863 /gene="nAChRalpha-30D" /note="ion channel; neurotransmitter transmembrane receptor; alternatively spliced; contains exons 3b and 8a" /codon_start=1 /product="nicotinic acetylcholine receptor Dalpha6 subunit variant type II" /protein_id="AA013393.1" /db_xref="GI:20152846" /translation="MDSPLASLSFVLLIFLAIKESQCPHEKRLNHLHLLSTNTL ERPVAESEPLEVKGFLTQOIIDYDEKNQLLITNLISLEWNLNLRNTEYGGVK DLRITPNKLWKPDVLMVANSADGFTYHTNIVKHGSGCLYVPGIFKSTCKMDITW PPDDOHCMEKFGSWTYDGNOLDVLNSEDGGLSDFTNGEWYLLAMPKKNITVYA CCPEPYVDITFTIQRRLTYFFNLIVPCVLISSMALLGFTLPDPSGKLTIGVTIL
gene	
CDS	

943	Db	GAAGATGAGGGGATCTTTCCGATTTTCATAACAAATGGCGAGTGGTACTTGTCTTGC	1002
491	Qy	ProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThr	510
1003	Db	CCGGGAAGAAGAATACGATAGTCTACGCTGCTGCCAGAACCATATGCGATATCACC	1062
511	Qy	PheAlaIleIleIleArgArgArgThrIleuTyrTyrPhePheAsnLeuIleIleProCys	530
1063	Db	TTTACTATACAAATTCGTCGCGGTACATPATATATTTTCAATTTAAATCGTGCCATGT	1122
531	Qy	ValLeuIleAlaSerMetAlaLeuLeuGlyPheThrIleuProProAspSerGlyGluLys	550
1123	Db	GTGCTAATCTCATCGATGCCCTTACTGGGCTCTCAATTCGCCCGCGAATTCGGCGGAGAA	1182
551	Qy	LeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGlu	570
1183	Db	CTGACGCTGGGAGTTACAAATTCCTTCTATCGCTCACAGTGTTCTCAACCTTGTAGCTGAG	1242
571	Qy	ThrMetProAlaThrSerAspAlaValProLeu	581
1243	Db	ACATTGCCCAAGTATCTGATGCAATCCCTTGTGTAGGCACCTACTTCAATTTGCATCATG	1302
581	Qy	-----	581
1303	Db	TTCATGTGCGCTCGTGGTGTGCTGACAGTAGTGTGTCTCACTACCACCATCGCAC	1362
582	Qy	-----TrrIleArgIleValPheLeuCysTrrLeuProTrr	593
1363	Db	CGCGACATTACAGAGATGCCACCGTGGATCAAGTCCGTTTTCTTACAAATGGCTGCCCTGG	1422
594	Qy	IleLeuArgMetSerArgProGlyArgProLeuIleLeuGluPheProThrThrProCys	613
1423	Db	ATCTTGCGAATGGGTGCGACCGGTGCGAAGATT-----	1455
614	Qy	SerAspThrSerSerGluArgGlyHisGlnIleLeuSer-----AspValGlu	629
1456	Db	-----ACACGCAAAACAATACTATTAAAGCAATCGCATGAAGGAGCTGGAG	1500
630	Qy	LeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspAspPhe	649
1501	Db	CTAAAGGAGCGCTCTCCAAATCCCTGCTGGCCAATGCTCTCGACATCGACGACGACTTC	1560
650	Qy	ArgHisAsnCysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyr	669
1561	Db	CGGCACACAATA-----	1572
670	Qy	ArgThrValTyrGlyGlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMet	689
1573	Db	-----TCTGCTCCCAACCGCATTTGGCTGTCTGGCCGAGCTTCGGTCGG	1617
690	Qy	ProAspAlaVal---ThrHisHisThrCysIleLysSerSerThrGluTyrGluLeuGly	708
1618	Db	CCCACAAACGGTGGAGGAGCATCACACGGCCATC---GGCTGCAATCACAAAGATCTTCAT	1674
709	Qy	LeuIleLeuLysGluIleArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsn	728
1675	Db	CTAATTCTCAAGAATTGCAATTTATTCGCGCGGATGGCCAAAGCTGACGACGAAGCG	1734
729	Qy	AspIleAlaAsnAspTrrLysPheAlaAlaMetValValAspArgLeuCysIleIleIle	748
1735	Db	GAATTGATCGCGGATTTGGAAGTTTCGCGCAATGGTTGTGGATAGATTTTGTTTAATGTT	1794
749	Qy	PheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSerAlaProHisIleIle	768
1795	Db	TTACGCTCTTCAGGATTTATGCAACGGTTACGGTGTCTCTCCGCTCGGCACATATC	1854
769	Qy	Val 769	
1855	Db	GTG 1857	
RESULT 11			
DME554209			
LOCUS			
		1665 bp	mRNA
		linear	INV 01-APR-2003

RESULT 11
DME554209
LOCUS

```
Qy 371 MetAsnLeuArgTTPAsnThrSerAspTyrGlyValVallysAspLeuArgIleProPro 390
Db 308 TACAAATCTGCGTGGATGAACCGAATACCGCGGGTCAAGGATCTACGATCACGCC 367
Qy 391 HisArgIleThrPlyProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGly 410
Db 368 AACAAAGTGTGGAAGCCGACGCTGCTCATGTACAAACAGCGCGGATGAGGATTCGATGCC 427
Qy 411 ThrTyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProGly 430
Db 428 ACGTATCACACGCGGTGTGGTCAACATCGCGCGAGTGTCTGTACGTGCCCGCTGGT 487
Qy 431 IlePheLysSerThrCysLysIleAspIleThrPheProPheAspAspGlnArgCys 450
Db 488 ATCTTCAAGAGCACATGCAAGATGACATACAGTGGTTCCTCATTTGATGACCAATGTC 547
Qy 451 GluMetLysPheGlySerThrTyrAspGlyPheGlnLeuAspLeuGlnAsp 470
Db 548 GAAATGAAATTCGGTAGTGGACTTACGATGGAATCAGTTGGATTTGGTTTGGATTGCC 607
Qy 471 GluThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTyrGluLeuGlyVal 490
Db 608 GAAGATGAGGAGGATCTTTCGATTTTCATAAATGGCAGTGGTACTTGTCTGGCCATG 667
Qy 491 ProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThr 510
Db 668 CCGGGAAGAGATACATAGTCTACGCGCTGCTGCCAGAACCATATGTCATATCACC 727
Qy 511 PheAlaIleIleArgArgThrLeuTyrTyrPhePheAsnLeuIleProCys 530
Db 728 TTTACTATACAAATTCGTCGCGTACATATATATATTTTCAATTTAATGTGGCATGT 787
Qy 531 ValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProAspSerGlyGluLys 550
Db 788 GTGCTAATCTCATGATGGCCCTACTGGGCTTTCACATTCGCGCGGATTCGGGGAGAAA 847
Qy 551 LeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGlu 570
Db 848 CTGACGCTGGCGTAACATATACTACTATCAATACAGATATTTCAACCTGTGCCGAG 907
Qy 571 ThrMetProAlaThrSerAspAlaValProLeu----- 581
Db 908 TCCATGCCGACACGTCGATGCTGTCTCTTATAGGACCTACTTCAATTTGCATCATG 967
Qy 581 ----- 581
Db 968 TTCATGTGCGCTCGTCGGTGGTGTGACAGTAGTGTGCTCAATACACATCGCAC 1027
Qy 582 ----- 582
Db 1028 GCGGACATTCACGAGATGCCACCGTGGATCAAGTCGTTTCTTCAATAGTGGTGGCTGG 1087
Qy 594 IleLeuArgMetSerArgProGlyArgProIleLeuLeuGluPheProThrThrProCys 613
Db 1088 ATCTTGCAGATGGTTCGACCGCGTGCAGAT----- 1120
Qy 614 SerAspThrSerSerGluArgGlyHisGlnIleLeuSer----- 629
Db 1121 ----- 1121
Qy 630 LeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPhe 649
Db 1166 CTAAAGGAGCGCTCTCCAAATCCCTGTGGCCATGCTCTGACATCGACATCGACGACTTC 1225
Qy 650 ArgHisAsnCysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyr 669
Db 1226 CGGCACACAATA----- 1237
Qy 670 ArgThrValTyrGlyGlnGlyAspGlySerIleGlyProIleGlySerThrArgMet 689
Db 1238 -----TCTGGCTCCCAACCGCCATTTGGTCTGTCGCCGACGCTTCGTCGG 1282
```

```
690 ProAspAlaVal---ThrHisHisThrCysIleLysSerSerThrGluTyrGluLeuGly 708
1283 CCCAACCGCTGGAGGAGCATCACACGCCCATC---GGCTGCAATCAAAAGATCTTCAT 1339
709 LeuIleLeuLysGluIleArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsn 728
1340 CTAATTCCTAAAGAAATTCGAATTTATTACGCGCGGATGCGCAAGCTGACGACGAACGC 1399
729 AspIleAlaAsnAspTrpLysPheAlaAlaMetValValAspArgLeuCysLeuIleIle 748
1400 GAATTGATCGCGATTCGAAGTTCGCGCAATGTTGGATAGATTTTGTAAATGTT 1459
749 PheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSerAlaProHisIleIle 768
1460 TTCACGCTCTTACGATTTATTGCAACGTTACGCTGCTCTCCGTCGCCACATAATC 1519
769 Val 769
1520 GTG 1522
```

RESULT 12
BT011147
LOCUS BT011147 1699 bp mRNA linear INV 19-DEC-2003
DEFINITION Drosophila melanogaster GH1518 full insert cDNA.
ACCESSION BT011147
VERSION BT011147.1 GI:40216007
KEYWORDS FLI CDNA.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
1 (bases 1 to 1699)
Stapleton, M., Brokstein, P., Hong, L., Agbavani, A., Carlson, J.,
Champe, M., Chavez, C., Dorsett, V., Dresnek, D., Farfan, D., Friese, E.,
George, R., Gonzalez, M., Guarin, H., Krommiller, B., Li, P., Liao, G.,
Miranda, A., Mungall, C. J., Nunoo, J., Paclebb, J., Paragas, V., Park, S.,
Patel, S., Phouanavong, S., Wan, K., Yu, C., Lewis, S. E., Rubin, G. M.
and Celisnik, S.
Direct Submission
Submitted (19-DEC-2003) Berkeley Drosophila Genome Project,
Lawrence Berkeley National Laboratory, One Cyclotron Road,
Berkeley, CA 94720, USA
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory
Berkeley, CA 94720
This clone was sequenced as part of a high-throughput process to
sequence clones from Drosophila Gene Collection 1 (Rubin et al.,
Science 2000). The sequence has been subjected to integrity checks
for sequence accuracy, presence of a polyA tail and contiguity
within 100 kb in the genome. Thus we believe the sequence to
reflect accurately this particular cDNA clone. However, there are
artifacts associated with the generation of cDNA clones that may
have not been detected in our initial analyses such as internal
priming, priming from contaminating genomic DNA, retained introns
due to reverse transcription of unspliced precursor RNAs, and
reverse transcriptase errors that result in single base changes.
For further information about this sequence, including its location
and relationship to other sequences, please visit our Web site
(http://fruitfly.berkeley.edu) or send email to
cdna@fruitfly.berkeley.edu.
location/Qualifiers
1. 1699
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/strain="y; cn bw sp"
/db_xref="taxon:7227"
<1. 1699
/gene="nAcRalpha-30D"
/db_xref="FLYBASE:FBgn0032151"
<1. 1528
/gene="nAcRalpha-30D"

FEATURES
source
mol_type="mRNA"
strain="y; cn bw sp"
db_xref="taxon:7227"
1. 1699
gene
/gene="nAcRalpha-30D"
db_xref="FLYBASE:FBgn0032151"
1. 1528
CDS
/gene="nAcRalpha-30D"

ORGANISM	Drosophila melanogaster	Db	406	TCGCTGTTTGTCTGTTGATCTTCTTGGCGATAATAAAGAAAGCTGT---CAAGGACCT	462
REFERENCE	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.	Qy	311	HisGluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProVal	330
AUTHORS	1 (bases 1 to 2023)	Db	463	CATGAAAGCGCTGCTGAACCATCTGCTGCCACCTACATACCTGGAGCGACCGCTG	522
TITLE	Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.	Qy	331	LeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleAsp	350
JOURNAL	Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalphas5, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing	Db	523	GCCAAATGAAATGAGAGCCCTGGAGTTAAGTTGGAGCTGACGCTGCAGCAGATCATGC	582
MEDLINE	Genetics 160 (4), 1519-1533 (2002)	Qy	351	ValAspGluLysAsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAsp	370
PUBMED	21969411	Db	583	GTGACGAGAGAAATCAGCTTCTCATACGAATCTTTGGCTTTTCGTGGAGTGGACGAC	642
REFERENCE	11973307	Qy	371	MetAsnLeuArgTrpAsnThrSerAspTyrGlyValLysAspLeuArgIlePro	390
AUTHORS	2 (bases 1 to 2023)	Db	643	TACAAATCTGCCTGGAATGAACCGAATACGCGGGGTCAAGGATCTACGAATCAGGCC	702
TITLE	Grauso, M. and Sattelle, D.B.	Qy	391	HisArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGly	410
JOURNAL	Direct Submission	Db	703	AACAAGCTGTGGAAGCCGACGCTCATGTACAACAGCGCGGATAGGGATTTCGATGCG	762
TITLE	Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional Genetics Unit, University of Oxford, South Park Road, Oxford OX1 3QX, UK	Qy	411	ThrTyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProProGly	430
FEATURES	Location/Qualifiers	Db	763	ACGTATCACACCAACATTTGCTGTCACCATAGCGGAGTTGTCTGTACGTGCCCCCTGGT	822
source	1..2023	Qy	431	IlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCys	450
gene	/organism="Drosophila melanogaster"	Db	823	ATCTTCAGACACATGCAAGATGGACATCACGTGGTTCCCATTTGATGACCAACATTGC	882
CDS	/mol_type="mRNA"	Qy	451	GluMetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnAsp	470
	/db_xref="taxon:7227"	Db	883	GAATGAAATTCGGTAGTTGGACTTACGATGAAATCAGTTGGATTGTTGTTGAATTC	942
	/chromosome="2"	Qy	471	GlnThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuGlyVal	490
	/map="30D1"	Db	943	GAAGATGGAGGGGATCTTTCCGATTTTCATAACAAATGGCGAGTGTACTTGTCTGCCCATG	1002
	/dev_stage="embryo"	Qy	491	ProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThr	510
	1..2023	Db	1003	CGGGAAGAAGAATACGATAGTCTACGCTGCTGCCAGAACCATATGTCGATATCAC	1062
	/gene="nAcRalpha-30D"	Qy	511	PheAlaIleIleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCys	530
	/note="ion channel: neurotransmitter transmembrane receptor; alternatively spliced; contains exons 3b and 8b"	Db	1063	TTTACTATACAAATTCGTCGCGTACATTATATTATTTTCAATTTAATCGTGCATGT	1122
	/codon_start=1	Qy	531	ValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLys	550
	/product="nicotinic acetylcholine receptor Dalphas6 subunit variant type III"	Db	1123	GTGCTTAATCTCATCGATGCGCTTACTGGGCTTCACATTGGCGCGGATTCGGCGGAGAAA	1182
	/protein_id="AA013394.1"	Qy	551	LeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGlu	570
	/db_xref="GI:20152849"	Db	1183	CTGAGCTGGGAGTTACAAATCTTCTATCGCTCACAGTGTTCCTCAACCTTGTAGTGAG	1242
	/translation="MDSPLPASLSLFLVLLIFLAIKESCGPHEKRLNHLSTVNTL ERVANSEPLEVFGTLQQLIDVEKNQLITNLWLSLEWNDNLKWNTEYGGVK DLRTTPKLMKPDVLMYNSADEGFDGTYTNIIVKHSGSLYVPPGIFKSTCKMDITW FPFDDHCKMFGSWTYDGNQLDLVNSEDDGSLDSFITNGEWLLAMPKGNKTIVYA CCPPYVDITFTIQRRTLYFFNLIVPCVLISMAILLGLTPDPSGEKLTIGVTIL LSLTVFLNLVETLPOVSDAIPLLGTNYFNCIMFMVASSVLTVVVNLVNHRTADITHM PFWKSVFLQMLPWLPMRGPRCKITRTKILLNEMKELEKRSKSLLANVLDDID DFRHTISGQTALGSSAFGRPTTVEEHTAIGCNKOLHLILKELQFITARMRKADD EAEILGDKWKFAMVVDVRFLLIVFTLITTIATVTLLSAPHIIQ"	Qy	571	ThrMetProAlaThrSerAspAlaValProLeu	581
variation	794	Db	1243	ACAAATGCCCAAGTATCTGATGCAATCCCTTTAGGCACCTACTTCAATTCATCATG	1302
	/note="results in glycine to serine substitution; compared to the sequence deposited in GenBank Accession Number AE003626"	Qy	581	-----	581
variation	846	Db	1303	TTCAATGGTGGCTCGTGGTGGTGTGACAGTAGTGGTGTCTCACTACCACTCGCACA	1362
	/replace="a"	Qy	582	-----	582
	/gene="nAcRalpha-30D"	Db	1363	CGGACATTCACGAGATGCCACCGTGGATCAAGTCGTTTCTCAATAATGGTGGCTGG	1422
	/note="results in asparagine to serine substitution; compared to the sequence deposited in GenBank Accession Number AE003626"	Qy	594	IleLeuArgMetSerArgProGlyArgProLeuIleLeuGluPheProThrThrProCys	613
	794	Db	1423	ATCTTGGAAATGGTTCGACCGCGTCGCAAGATT	1455
	/note="results in asparagine to serine substitution; compared to the sequence deposited in GenBank Accession Number AE003626"	Qy	614	SerAspThrSerSerGluArgLysHisGlnIleLeuSer	629
	794	Db	1456	-----	1456
	/replace="a"	Qy	291	SerTrpIlePheLeuIleTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyr	310
ORIGIN	US-09-303-232-2 (1-770) x AF321447 (1-2023)	Db			
Alignment Scores:					
Pred. No.:	5,94e-109				
Score:	1620.50	Length:	2023		
Percent Similarity:	70.44%	Matches:	318		
Best Local Similarity:	61.04%	Conservative:	75		
Query Match:	40.08%	Mismatches:	79		
DB:	3	Indels:	7		
		Gaps:	7		

```

Qy 630 LeuLysGluArgSerSerLysSerLeuLeuAlaAsnValLeuAspLeuAspPhe 649
Db 1501 CTAAGAGCGCTCTCTCAAAATCCCTGCTGCCAATGTCCTCGACATCGACGACACTTC 1560
Qy 650 ArgHisAsnCysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyr 669
Db 1561 CGGCACACAATA----- 1572
Qy 670 ArgThrValTyrGlyGlnGlyAspAspCysSerIleGlyProIleGlySerThrArgMet 689
Db 1573 -----TCTGGCTCCCAACCCCAATGCTGCTCGGCAGCTTCGGTCGG 1617
Qy 690 ProAspAlaVal-----ThrHisHisThrCysIleLysSerSerThrGluTyrGluLeuGly 708
Db 1618 CCCACACGGTGGAGGACCATCACACGCCCATC---GGCTCAATCAAAAGATCTTCAT 1674
Qy 709 LeuIleLeuLysGluIleArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsn 728
Db 1675 CTAATTCCTCAAGAAATGCAATTTATTACGGCGCGGATGCGCAAGCTGACGACGAAGCG 1734
Qy 729 AspIleAlaAsnAspTyrLysPheAlaMetValValAspArgLeuCysLeuIleIle 748
Db 1735 GAATTTGATCGCGGATGGAAGTTCGCGGCAATGTTGTGGATAGATTTTGTAAATGTT 1794
Qy 749 PheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSerAlaProHisIleIle 768
Db 1795 TTCACGCTCTTCAGATTAATGCAACGGTACGGTGTGCTCTCGCTCCGCATCAATC 1854
Qy 769 Val 769
Db 1855 GTG 1857

RESULT 14
LOCUS AF321449
DEFINITION Drosophila melanogaster 2110 bp mRNA linear INV 29-APR-2002
subunit variant type V (nAcRalpha-30D) mRNA, complete cds,
alternatively spliced.
ACCESSION AF321449
VERSION AF321449.1 GI:20152852
KEYWORDS
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Phytophthora; Drosophilidae; Drosophila.
1 (bases 1 to 2110)
Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalphea5, Dalphea6 and Dalphea7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
Genetics 160 (4), 1519-1533 (2002)
21969411
11973307
REFERENCE 2 (bases 1 to 2110)
AUTHORS Grauso, M. and Sattelle, D.B.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
Genetics Unit, University of Oxford, South Park Road, Oxford OX1
3QX, UK
FEATURES
source
1..2110
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="30D1"
/dev_stage="embryo"
1..2110
/gene="nAcRalpha-30D"
379..1950
CDS

```

```

/gene="nAcRalpha-30D"
/note="ion channel; neurotransmitter transmembrane
receptor; alternatively spliced; contains exons 3a, 8a and
8b; contains two repeats of the transmembrane TM2 region"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalphea6 subunit
variant type V"
/protein_id="AAM13396.1"
/db_xref="GI:20152853"
/translation="MDSPLPASISLFVLLIFLAIKESQCPHEKRLINHLSTYNTL
ERPVAESEPLEVKFGLTLOOIIDVDEKNOILITNAMLEWNYDLNRWNETYGVK
DLRTTPNKLKPDVLMYNSADEGDTYHTNIVVKNHSGCLYVPGGFKSTCKMDITW
FPDDDOCEMKFGSWYDGNQLDLVNSEGGDLSDFITNGEWLLAMPKKNITVIA
CCPEPYVDITFTTQIRRTTYFFNLIVPCVLISSMALLGFTIPDPDSGEKLTGLVTIL
LSUTFVNLVAESMPTSDAVPLIGVTLLSLTVFLNLVAETLPQVSDAIPLLGTYN
CIMFVASSVLTVVNLVHRTADIHMPFKSVPLQWLPLMRGRPKTRKT
IILSNRMKELEKERSKLLANVLDDDDFRHTISGQTAIGSSASFGPTTVEREH
TAIGCNKHDLHLKELQFITARMRRADDEALLIGDKFAAMVVDFCLVFTLPTIL
ATVTLLSAHHIIVQ"
846
/variation
/gene="nAcRalpha-30D"
/note="results in methionine to isoleucine substitution;
compared to the sequence deposited in GenBank Accession
Number AE003626"
/replace="a"
1401
/variation
/gene="nAcRalpha-30D"
/note="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replace="c"
1647
/variation
/gene="nAcRalpha-30D"
/note="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replace="c"
1770
/variation
/gene="nAcRalpha-30D"
/note="compared to the sequence deposited in GenBank
Accession Number AE003626"
/replace="c"

ORIGIN
Alignment Scores:
Pred. No.: 3,11e-108 Length: 2110
Score: 1611.00 Matches: 320
Percent Similarity: 66.55% Conservative: 46
Best Local Similarity: 58.18% Mismatches: 76
Query Match: 39.85% Indels: 108
DB: 3 Gaps: 7

US-09-303-232-2 (1-770) x AF321449 (1-2110)
Qy 291 SerTrpIlePheLeuLeuIleTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyr 310
Db 406 TCGCTGTTTCTCTCTGTTGATCTTTCTGGCATATTAAGAAAGCTGT---CAAGACCT 462
Qy 311 HisGluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProVal 330
Db 463 CATGAAAAGCGCTGCTGAACCATCTGCTCCACCTACAAATACGCTGGAGCGACCGTG 522
Qy 331 LeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleIleAsp 350
Db 523 GCCAATGAATCGAGCGCCCTGGAGGTTAAGTTGGACTGACGCTGCAGGAGATCATCGAC 582
Qy 351 ValAspGluLysAsnGlnLeuLeuValThrAsnValTyrLeuLysLeuGluTyrAsnAsp 370
Db 583 GTGGATGAAAAGATCAGATTCTGACCAAAATCGCTGGTTAAATTTGAGTGGACAGAC 642
Qy 371 MetAsnLeuArgTrpAsnThrSerAspTyrGlyValLysAspLeuArgIleProPro 390
Db 643 TACAATCTCGCTGGAATGAAACGGAATACGCGGGGTCAAGGATCTACGAATCACGCC 702
Qy 391 HisArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGly 410

```

```

Db 703 AACAGCTGTGGAGCCCGAGTCTCATGTACACGCGGATGAGGATTCGATGC 762
Qy 411 ThrTyrGlnThrAsnValValArgAsnGlySerCysLeuTyrValProGly 430
Db 763 ACGTATCACACCAATGTGTCAAAACAAACGCGAGTTGTCTGTACGTGCCCCCTGGT 822
Qy 431 IlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCys 450
Db 823 ATCTTCAAGAGCACATGCAAGATGGACATCACGTGGTTCCTCATTTGATGACCAATTC 882
Qy 451 GluMetLysPheGlySerThrTyrAspGlyPheGlnLeuAspLeuGlnAsp 470
Db 883 GAATGAATTCGAGTGTGACATACGATGAATCAGTTGATGTTTGAATTC 942
Qy 471 GluThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTyrGluLeuGlyVal 490
Db 943 GAAGATGGAGGGATCTTTCGATTTTCAACAAATGCGAGTGTGATCTTGTTCGCCATG 1002
Qy 491 ProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThr 510
Db 1003 CCGGAAAGAGATAGATAGTCTACGCTGCTGCCAGAACCATATGTCGATATACC 1062
Qy 511 PheAlaIleIleIleArgArgThrLeuTyrTyrPheAsnLeuIleIleProCys 530
Db 1063 TTTACTATACAAATTCGTCGCGTACATATATATTTTCAATTTAATCGTCCCATGT 1122
Qy 531 ValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLys 550
Db 1123 GTGCTAATCATCATGATGGCCCTACTGGGCTTCATTTGCCCGCGGATTCGGCGAGAA 1182
Qy 551 LeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGlu 570
Db 1183 CTGACGCTGGCGTAACTACTACTATCATTAACAGTATTTCTAAACCTTGTGCCGAG 1242
Qy 571 ThrMetProAlaThrSerAspAlaValProLeu----- 581
Db 1243 TCCATGCCGACACGTCGGATGCTTCTCTTATAGGATTAACAATCTTCTATCGCTC 1302
Qy 581 ----- 581
Db 1303 ACAGTGTCTTCAACCTTTAGCTAGACATTCGCCCAAGTATCTGATGCAATCCCTTG 1362
Qy 581 ----- 581
Db 1363 TTAGCACCTACTCAATTCATGTCATGTTCTATGTCGCATCGTCGGTGGCTGACAGTA 1422
Qy 582 ----- 582
Db 1423 GTGGTGTCAACTACCATCGCACAGCGGACATTCACGAGATGCCACCGTGGATCAAG 1482
Qy 585 IleValPheLeuCysTrpLeuProTrpIleLeuArgMetSerArgProGlyArgProLeu 604
Db 1483 TCCGTTTTCTTCAACATGCTCCCTCGGATCTTGGCAATGGTTCGACCGCGTCGAAGAT 1542
Qy 605 IleLeuGluPheProThrThrProCysSerAspThrSerSerGluArgLysHisGlnIle 624
Db 1543 ----- 624
Qy 625 LeuSer-----AspValGluLeuLysGluArgSerSerLysSerLeuLeuAla 640
Db 1561 TTAAGCAATCGCATGAAGGAGCTGGAGCTAAAGGAGCGCTCTCCAAATCCCTGTCGCG 1620
Qy 641 AsnValLeuAspIleAspAspPheArgHisAsnCysArgProMetThrProGlyGly 660
Db 1621 AATGTCCTCGACATCGACGACACTTTCGGCACACAATA----- 1659
Qy 661 ThrLeuProHisAsnProAlaPheTyrArgThrValTyrGlyGlnGlyAspAspGlySer 680
Db 1660 ----- 680
Qy 681 IleGlyProIleGlySerThrArgMetProAspAlaVal-----ThrHisHisThrCysIle 699

```

```

Db 1678 ATTGGCTCGTCGCGCAGCTTCGGTCGCCCCACACACGCTGGAGGAGCATCACACGCCCATC 1737
Qy 700 LysSerSerThrGluTyrGluLeuGlyLeuIleLeuLysGluIleArgPheIleThrAsp 719
Db 1738 ---GGTGCATACAAAGATCTTCATCTAATTTCTTAAGAATTCGAATTTATTACGGCG 1794
Qy 720 GlnLeuArgLysAspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMet 739
Db 1795 CGGATCGCAAAAGCTGACGACGAGCGAATTCGTCGCGATTCGGAAGTTTCGCGCAATG 1854
Qy 740 ValValAspArgLeuCysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAla 759
Db 1855 GTTGTGATAGATTTTGTAAATTTTTCACGCTCTTCAGATATTATTCGAACGGTTACG 1914
Qy 760 ValLeuLeuSerAlaProHisIleIleVal 769
Db 1915 GTGCTGCTCTCCGTCGCCACATAATCGTG 1944

RESULT 15
AF143847 3029 bp mRNA linear INV 27-MAY-1999
LOCUS Heliothis virescens putative nicotinic acetylcholine receptor alpha
DEFINITION 7-2 subunit mRNA, complete cds.
ACCESSION AF143847
VERSION AF143847.1 GI:4895006
KEYWORDS Heliothis virescens (tobacco budworm)
SOURCE Heliothis virescens
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1 (bases 1 to 3029)
AUTHORS Schulte, T., Oellers, N. and Adamczewski, M.
TITLE Putative alpha subunits of insect nicotinic acetylcholine receptors
more similar to vertebrate alpha 7 subunits and C. elegans Ce21
than to other insect nicotinic acetylcholine receptor alpha
subunits
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3029)
AUTHORS Schulte, T., Oellers, N. and Adamczewski, M.
TITLE Direct Submission
JOURNAL Submitted (19-APR-1999) ZF-BTB, Bayer AG, Bldg. Q 18,, Leverkusen
51368, Germany

FEATURES
Location/Qualifiers
source 1..3029
/organism="Heliothis virescens"
/mol_type="mRNA"
/db_xref="taxon:7102"
95..1600
/feature="hvnachra7-2"
/codon_start=1
/product="putative nicotinic acetylcholine receptor alpha
7-2 subunit"
/protein_id="AAD32698.1"
/db_xref="GI:4895007"

CDS
/translation="MAPMALALALLPVSEQPEHEKRLNALLANNYNTLERPVANE
SELEVRFGTLQOIIDVENKQLITNLSLEWYDNLKWNDSYGVKDLRITPN
KLWKPDLVLYNSADEGDTGYQTVNVRSGGSLYVPGIFKSTCKMDIAWPFDDQH
CDMKPGSWTDGNDLVLKDEAGDLSDFITNGEWYLGMPGKNTITYACCPYV
DVFTETIRRTLYFFNLIVPCVLISSMALLGVTLVPPDSSEKLTGLVTLISLTVFL
NLVATETLPQVSDAIPLLGTIVFNCIMFMVAVSVLTVVNLVYHRTADIEHPQWIKSV
FLOWPLWILKMSRFGKTRKTMNTRMREKLKRSKSLANVLIDDDFRHGP
PPNSTAGTGLNGPSCIPTDFRFRSPRPTMEDVGGGLGSHHRELHLILELOPITA
RMKKADEAEELISDWKFAAMVVDVDFCLFVFTLTFTIATVAVLLSAPHIIVQ"

ORIGIN
Alignment Scores:
Pred. No.: 7.01e-108 Length: 3029
Score: 1609.00 Matches: 319
Percent Similarity: 71.62% Conservative: 57
Best Local Similarity: 60.76% Mismatches: 65
Query Match: 39.80% Indels: 84
DB: 3 Gaps: 10

```

US-09-303-232-2 (1-770) x AF143847 (1-3029)

Qy	295	LeuLeuIleTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyrHisGluLysArg	314	Db	1130	TCGAGGCCAGGAAGAAGATCACCAGGAAG-----ACTATAATGATGAACACGAGG	1180
Db	122	CTGTGGCTTTGGTGGCCGTATCGGACAA-----GGTCCTCACGAGAAGAGA	169	Qy	618	SerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGluArgSerSerLysSer	637
Qy	315	LeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeuAsnGluSer	334	Db	1181	-----ATGAGGAGCTGGAACTGAAGGAGAGGTCTCGAAGTCC	1219
Db	170	CTCCTGAACGGCTTGGCGGAACCTACAACACCCTGGAGCACCCTGGCGCAACGAGAGC	229	Qy	638	LeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsnCysArgProMetThr	657
Qy	335	AspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAspValAspGluLys	354	Db	1220	TTGTGGCGAATGTTCTAGATATTGATGACTTCAGACACGGC-----CCTCCGCCT	1273
Db	230	GAACCGCTAGAGTCAGGTCGGCTTGACCTTGACGAAATCATGTAGCTGGACGAGAAG	289	Qy	658	ProGly-----GlyThrLeuProHisAsnProAlaPheTyrArgThrVal	672
Qy	355	AsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArg	374	Db	1274	CTTAACAGTACTGCTCGACCGGAATTGGACCTGGGTGCTCAATATTCGACCGGAT	1333
Db	290	AATCAACTACTATAACCAATATATGCTGCTGGAGTGGATGAACTACTACAACCTGAGG	349	Qy	673	TyrGlyGlnGly-----AspAspGlySerIleGlyProIle	684
Qy	375	TrpAsnThrSerAspTyrGlyGlyValIysAspLeuArgIleProProHisArgIleTrp	394	Db	1334	TTCCGTGGGTGGTTCGTCGTCCTCCACGATGGAAGACGTGGGC-----GGCGGGCTG	1387
Db	350	TGGMACCAGCAGTATGGCGGGTCAAGACCTCAGGATCACGCCCAACAACTTGTGG	409	Qy	685	GlySerThrArgMetProAspAlaValThrHisHisThrCysIleLysSerSerThrGlu	704
Qy	395	LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThr	414	Db	1388	GGTAGC-----CACCATCGC-----	1402
Db	410	AAGCCGACGCTCTTATGTATATATAGTGTGACGAGGGTTTTCAGCGGACCTACCAACC	469	Qy	705	TyrGluLeuGlyLeuIleLeuLysGluIleArgPheIleThrAspGlnLeuArgLysAsp	724
Qy	415	AsnValValValArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSer	434	Db	1403	---GAGCTGCACCTCATCTAGAGAGCTGCAGTTCTACCGCCAGATGAAGAAGCT	1459
Db	470	AACGTGTGTGTCAAGACGGCGGAGTTGCTGTACGTGGCCACCTGGCATATTCAGAGC	529	Qy	725	AspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetValValAspArgLeu	744
Qy	435	ThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPhe	454	Db	1460	GATGAGGAAGCCGAGCTGATCAGCGACTGGAAAGTTTCTCGCATGGTTGTGTAGAGTTT	1519
Db	530	ACATGCAAGATGACATCGCGTGTGTTCCCTTCACCACCAACAACACTGTGATATGAAGTC	589	Qy	745	CysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSerAla	764
Qy	455	GlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAspGluThrGlyGly	474	Db	1520	TGCTGTTCGTGTACACTTTTCACAATCATCGCACAGTAGCTGTCTGTATCGGCA	1579
Db	590	GGTAGCTGGACATATGACGCAATCAGTTGGATCTGGTGTCTAAGATGAGGAGCGGC	649	Qy	765	ProHisIleIleVal	769
Qy	475	AspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLysArg	494	Db	1580	CCGATATATCATCTGTG	1594
Db	650	GATCTATCGGACTTCATAACAAATGGGAGTGGTATCTAATAGGAATGCCAGGCAAAAG	709				
Qy	495	AsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaIleIle	514				
Db	710	AAACAATACATACGGCTGCTGCCCGACCCCTACGTGGAGCTCACCTTCACCATCATG	769				
Qy	515	IleArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysValLeuIleAla	534				
Db	770	ATAAGAAGACGAACCTTGTTACTTCTTCAACCTGATCGTCCGTCGTGCTGAICTCA	829				
Qy	535	SerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGly	554				
Db	830	TCATGGCACTTCCTCGGCTTTCACACTGCCACACAGACTCCGGAGAGAACTCACACTTGA	889				
Qy	555	ValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAla	574				
Db	890	GTCACTATTCTTCTATCGCTGACGGTGTCTCAACCTGTGTAGCCGAGACCTGTGCCACAG	949				
Qy	575	ThrSerAspAlaValProLeu-----	581				
Db	950	GTCTCCGACGCTATCCCTCTTTAGGAGCGTACTTCAATTGTCATGTTTCATGGTAGCG	1009				
Qy	581	-----	581				
Db	1010	TCGTCTGTGTACTGACTGTGGTGGTACTCAATTAACCACTCGAACAGCTGATATACAT	1069				
Qy	582	-----	582				
Db	1070	GAATGCCACAGTGGATTAATCAGTATTCTCAATGGTGGCATGATGATCGCAATG	1129				
Qy	598	SerArgProGlyArgProLeuIleLeuGluPheProThrThrProCysSerAspThrSer	617				

Search completed: May 9, 2004, 00:54:29
Job time : 8539.79 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 06:39:04 ; Search time 808.783 Seconds
(without alignments)
4044.486 Million cell updates/sec

Title: US-09-303-232-2

Perfect score: 4043

Sequence: 1 MKNAQLKLTVDDELWLV.....MFAILATIAVLVSAPHIIVS 770

Scoring table:

BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:

-MODEL=frame+p2n.model -DEV=xlh
-Q=/cgn2_1/USPTO_spool/US09303232/runat_07052004_101109_23873/app_query.fasta_1.2261
-DB=N_geneseq_29Jan04 -QFW=fastap -SUFFIX=ring -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blom62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THRE SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFWT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09303232 @CGN 1.1200 @runat_07052004_101109_23873 -NCFU=6 -ICPU=3
-NO_WMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : N Geneseq_29Jan04.*
1: Geneseqn1980s.*
2: Geneseqn1990s.*
3: Geneseqn2000s.*
4: Geneseqn2001as.*
5: Geneseqn2001bs.*
6: Geneseqn2002s.*
7: Geneseqn2003as.*
8: Geneseqn2003bs.*
9: Geneseqn2003cs.*
10: Geneseqn2004s.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	4043	100.0	2886	3	Aaz24475
2	1850	45.8	3700	3	Aaz24476
3	1609	39.8	3109	3	Aaz24477
4	1498.5	37.1	984	4	ABL27131
5	1466.5	36.3	1540	4	ABL07231
6	1456.5	36.0	3144	4	ABL27130
7	1407.5	34.8	936	4	ABL13733
8	1253	31.0	803	4	ABL07799

9	1179.5	29.2	2769	2	AAT59196	Neuronal
10	1170.5	29.0	2101	2	AAT59197	Neuronal
11	1159.5	28.7	1509	3	AAC58395	Human PRO
12	1159.5	28.7	1509	4	AAC90380	Wild-type
13	1159.5	28.7	1876	2	AAV12197	Human neu
14	1159.5	28.7	1876	2	AAV12197	Human neu
15	1159.5	28.7	1876	6	ABS54875	Neuronal
16	1159.5	28.7	1876	6	ABS54875	Human neu
17	1159.5	28.7	1876	8	ADA10864	Human neu
18	1155.5	28.6	1590	2	AAV44687	V274T var
19	1153.5	28.5	1509	4	AAC90385	Human neu
20	1149.5	28.4	1509	4	AAC90386	Mutant hu
21	1143.5	28.3	1509	4	AAC90387	Mutant hu
22	1129	27.9	1964	6	ABZ11298	Human pol
23	1015.5	25.1	5096	4	ABL13732	Human pol
24	973	24.1	1915	4	AD20962	Caenorhab
25	935	23.1	1416	4	AAC90382	Chimeric
26	848.5	21.0	2277	2	AAV12199	Human neu
27	848.5	21.0	2277	2	AAQ90387	Human neu
28	848.5	21.0	2277	6	ABS54870	Human neu
29	848.5	21.0	2277	6	ABV73243	Human neu
30	848.5	21.0	2664	2	AAT48235	Neuronal
31	848.5	21.0	2664	6	ABK92165	Prostate
32	848.5	21.0	2664	8	ADA10854	Human neu
33	848.5	21.0	2664	9	ADC71170	Human neu
34	848.5	21.0	2666	9	ADC71168	Human 205
35	847.5	21.0	2319	4	ABL04391	Human 205
36	845	20.9	1908	2	AAT48236	Neuronal
37	845	20.9	1908	8	ADA10856	Human neu
38	844.5	20.9	2210	4	ABL10275	Drosophil
39	841.5	20.8	14668	4	ABL07798	Drosophil
40	841	20.8	2398	4	ABL26491	Drosophil
41	840	20.8	1584	7	ADA83809	Human CHR
42	834	20.6	1896	6	AAL45867	Modified
43	831.5	20.6	2385	4	ABL11821	Drosophil
44	829.5	20.5	1503	4	AAD20961	Caenorhab
45	826.5	20.4	2170	4	ABL09701	Drosophil

ALIGNMENTS

RESULT 1

Aaz24475

ID Aaz24475 standard; cDNA to mRNA; 2886 BP.

XX AC Aaz24475;

XX AC Aaz24475;

DT 17-FEB-2000 (first entry)

XX DE D. melanogaster acetyl-choline receptor DNA from clone Da7.

XX KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;

XX KW neurotransmission; plant protection agent; conductance; ACHR; ds.

XX OS Drosophila melanogaster.

XX FH Key

XX CDS Location/Qualifiers

FT 372..2684

FT /*tag= a

FT /product= "acetyl choline receptor."

XX PN DE19819829-Al.

XX XX 11-NOV-1999.

XX PF 04-MAY-1998; 98DE-01019829.

XX PR 04-MAY-1998; 98DE-01019829.

XX PA (FARB) BAYER AG.

XX PI Adamczewski M, Oellers N, Schulte T;

DR WPI: 2000-014207/02.
 DR P-PSDB; AAY50814.
 XX New nucleic acid encoding a nicotinic acetylcholine receptor from
 PT insects, used to identify potential insecticides.
 PT
 XX
 XX
 PS Claim 1a; Page 8-12; 26pp; German.

XX This invention describes a novel nucleic acid (NA) encoding a nicotinic
 CC acetyl-choline receptor (I) from insects which can be used as an
 CC insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
 CC (also vectors containing it, its regulatory regions, and antibodies
 CC directed against (I)-encoded proteins) are used to screen for: (a) plant
 CC protection agents that alter conductance of AChR, potentially useful as
 CC insecticides, or (b) genes which encode polypeptides that are involved in
 CC formation of functionally related AChR in insects. (I) are also used to
 CC isolate and characterize the specified regulatory regions and for
 CC recombinant production of (II). This sequence encodes an acetyl-choline
 CC receptor isolated from *Drosophila melanogaster*
 XX

SQ Sequence 2886 BP; 942 A; 687 C; 640 G; 617 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.:	0	Length:	2886
Score:	4043.00	Matches:	770
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	100.00%	Indels:	0
DB:	3	Gaps:	0

US-09-303-232-2 (1-770) x AAZ24475 (1-2886)

QY	1	MetLysAsnAlaGlnLeuLysLeuThrGluValAspAspGluLeuTrpLeuAlaVal	20
DB	372	ATGAAATGACAACTGAACTGACTGAGTTGACGATGATGAGCTGTGGTCGACGTA	431
QY	21	ArgLeuAlaHisCysSerSerAsnPheSerSerSerSerThrArgThrThrSerSer	40
DB	432	AGATTAGCGCACTGCAGCAGCAACTTTAGCAGCAGTAGCAGCAGCAAGAACCCACGACG	491
QY	41	AsnGlnArgHisAsnGlnGlnLeuThrThrLeuGlnProArgSerLeuSerThrLysHis	60
DB	492	AACGAGAGGCAACACGACCACTCAACACTGCAACCAAGAGAGCTTAAGTACAAAAAC	551
QY	61	HisSerAsnIleAlaSerGluGlnHisAsnSerGlnGlnGlnGluProAlaSerLysAsp	80
DB	552	CACAGCAACATTGCAAGCGAGCAGCAGCAATAGCCAGCAACAGGAGCCAGCATCGAAGGAC	611
QY	81	GluAspValAlaAsnHisGlyArgSerAsnAspGlnGlnThrHisLeuGlnGlnLeuAsp	100
DB	612	GAGGATGTAGCCACCAACCGTAGAAGCAATGACCAAGCAGCAGCAGCATCTGCACAGCTAGAC	671
QY	101	SerSerAsnMetLeuSerProLysThrAlaAlaAlaAlaThrAlaAlaGlyAspGluAla	120
DB	672	AGCAGCAACATGTTGTCGCCAAGACAGCGCAGCAGCAACTGTCGCCGCGGATGAAGCA	731
QY	121	ThrThrGlnGlnProThrAsnIleArgLeuCysAlaArgLysArgGlnArgLeuArgArg	140
DB	732	ACAAACCAACCAACCAACCAATAGACTGTGTGACGCAACGAGCAACGATGTGTCGCG	791
QY	141	ArgArgLysArgLysProAlaThrProAsnGluThrAspIleLysLysGlnGlnGlnLeu	160
DB	792	CGACGAAAAAGAAACACGCAACCCCAACAGAAACAGATATCAAGAAACCAACAGCAACTT	851
QY	161	SerMetProProPheLysThrArgLysSerThrAspThrThrThrThrProAlaThr	180
DB	852	AGCATGCTCCCTTCAAAACGCGCAATCCAGCACACCTACAGCACACCAAGCAACA	911
QY	181	ThrSerCysProThrAlaThrThrMetGlnCysArgAlaSerAspAsnGluPheSerIle	200
DB	912	ACAGCTGTCGACAGCAACCTACATGCAATGTGACGACGACGACCAATGAGTTTCAGTAT	971
QY	201	ProLysSerArgHisAspArgValSerThrAlaThrPheAlaThrValLeuHisValLeu	220

DB	972	CCGATATCGAGACATGATAGATTATCCACGCGCACATTCGCTGGTGTTCATGTGCTG	1031
QY	221	GlnValLeuLeuValSerLeuGlnGlnTrpGlnLeuHisValGlnGlnArgSerValLeu	240
DB	1032	CAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1091
QY	241	LeuPheArgArgIleAlaAlaSerThrIleAlaPheIleSerTyrLeuGlySerPheAla	260
DB	1092	CTGTTTCAAGAGATCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1151
QY	261	AlaGlnLeuLysAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer	280
DB	1152	CGCAACTGAAAAATAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1211
QY	281	ThrGlnIleLeuAsnGlyLeuAsnLysHisSerTrpIlePheLeuLeuIleTyrLeuAsn	300
DB	1212	ACGCAATATTTAAACGGACTTTAAACACTCATGGATATTTTATTGATATATTTGAT	1271
QY	301	LeuSerAlaLysValCysLeuAlaGlyTyrHisGlyLysArgLeuLeuHisAspLeuLeu	320
DB	1272	TTATCTGCTAAAGTTTCTAGCAGGATATCATGAAAGAGACTGTTACACGATCTTTTG	1331
QY	321	AspProTyrAsnThrLeuGluArgProValLeuAsnGluSerAspProLeuGlnLeuSer	340
DB	1332	GATCCTTATAATACACTAGAACGTCCTGCTCAATGAATCGGACCCGCTTACAATTAGC	1391
QY	341	PheGlyLeuThrLeuMetGlnIleAspValAspGluLysAsnGlnLeuLeuValThr	360
DB	1392	TTTGGTTTAACTTTAATGCAAAATATCGATGTGACGAGAAATCAATTCCTAGTCACT	1451
QY	361	AsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspTyr	380
DB	1452	AATGTGCTTAAACTGGAGTGGACGACATGATCTCCGCTGGAACACCTCCGACTAT	1511
QY	381	GlyGlyValLysAspLeuArgIleProProHisArgIleTrpLysProAspValLeuMet	400
DB	1512	GGCGAGTTAAGGATCTGCGAATACCGCGCATCGCATCTGGAAGCCGAGCTGCTGATG	1571
QY	401	TyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThrAsnValValValArgAsn	420
DB	1572	TACAAAGTGGGATGAGGATTTGACGGACCTTACAGAGCAAGCTGCTGTCGGAAC	1631
QY	421	AsnGlySerCysLeuTyrValProProGlyIlePheLysSerThrCysLysIleAspIle	440
DB	1632	AACGCTCGTCTATACGTTCCGCGGGATCTTCAAGTCGACGTCGACGATCGACATC	1691
QY	441	ThrTrpPheProPheAspGlnArgCysGluMetLysPheGlySerTrpThrTyrAsp	460
DB	1692	ACGTGGTTCCCTTCGATGACGAGCGGTGCGAGATGAAGTTCCGCGAGTTGGACCTTACC	1751
QY	461	GlyPheGlnLeuAspLeuGlnLeuGlnAspGluThrGlyGlyAspIleSerSerTyrVal	480
DB	1752	GGATTCAGCTGGATTTACAAATTACAGATGAATCGCGGTGATATCAGCAGTTACGCTG	1811
QY	481	LeuAsnGlyGluTrpGluLeuGlyValProGlyLysArgAsnGluIleTyrTyrAsn	500
DB	1812	CTCAACGCGAGTGGAACTACTGGGTGTGCGGCAAAACGTAACAGATCTATTACAAC	1871
QY	501	CysCysProGluProTyrIleAspIleThrPheAlaIleIleArgArgArgThrLeu	520
DB	1872	TGCTGCCCGGAAACCCATATATAGACATACCTTCGCCATCATCTCCCGCAGGAACACTG	1931
QY	521	TyrTyrPhePheAsnLeuIleIleProCysValLeuLeuAlaSerMetAlaLeuLeuGly	540
DB	1932	TACTATTTCTCAACCTGATCATACCTTGTGTACTGATGCTCCATGCTGCTCGGA	1991
QY	541	PheThrLeuProProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeuLeuSer	560
DB	1992	TTACCCCTGCCCGCAGATTCGGGTGAAAAATATCGCTGGGTGTTACCATCTTGCTCTCG	2051
QY	561	LeuThrValPheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValPro	580

Db 2052 CTGACCGGTGTTCTGAATATGTTGCCGAGACAATGCCGGCTACTTCGGATCGGTGCCA 2111
 QY 581 LeuTrpIleArgIleValPheLeuCysTrpLeuProTrpIleLeuArgMetSerArgPro 600
 Db 2112 TTGTGGATACGATCGTGTGTTGTGCTGGCTGCCATGATATCGGAATGAGTCGCCCA 2171
 QY 601 GlyArgProLeuIleLeuGluPheProThrThrProCysSerSerAspThrSerSerGluArg 620
 Db 2172 GGACGACCGCTGATCTAGATGTTCCGACCCAGCCCTGTTCGGACACATCTCCGAGCGG 2231
 QY 621 LysHisGlnIleLeuSerAspValGluLeuLysGluArgSerSerLysSerLeuLeuAla 640
 Db 2232 AAGCACCAGATCTCCGACGTTGAGCTGAAAGAGCGCTCGTGAATCGTGTGCGC 2291
 QY 641 AsnValLeuAspIleAspAspPheArgHisAsnCysArgProMetThrProGlyGly 660
 Db 2292 AAGGTACTAGACATCGATGATGACTTCGGGCACAAATTCGCCCCCATGACGCCCGCGGA 2351
 QY 661 ThrLeuProHisAsnProAlaPheThrArgThrValTyrGlyGlnGlyAspAspGlySer 680
 Db 2352 ACATGTCACACAAACCCGCTTCTATCGCAGCGTTTATGGACAGGCGACATGGCGAGC 2411
 QY 681 IleGlyProIleGlySerThrArgMetProAspAlaValThrHisHisThrCysIleLys 700
 Db 2412 ATTGGGCCAATTGGCAGCACCCGATGCGGATGCGGTCAACCATCATACGTGCATCAA 2471
 QY 701 SerSerThrGluTyrGluLeuGlyLeuIleLeuLysGluIleArgPheIleThrAspGln 720
 Db 2472 TCATCACTGAATGATGAATAGGTTTAACTTAAAGGAAATTCGCTTTTATAACTGATCAG 2531
 QY 721 LeuArgLysAspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetVal 740
 Db 2532 CTACGTAAGATGACGAGTGCAATGACATTCGCAATGATTCGCAATTTGCGAGTATGGTC 2591
 QY 741 ValAspArgLeuCysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaVal 760
 Db 2592 GTTGACAGACTGTGCCTTATCATATTCACAATGTTGCAATATTAGCCACAATAGCTGTA 2651
 QY 761 LeuLeuSerAlaProHisIleIleValSer 770
 Db 2652 CTACTATCGGCACCATATATTATTGTCTCG 2681

RESULT 2

AAZ24476
 ID AAZ24476 standard; cDNA to mRNA; 3700 BP.
 XX
 AC AAZ24476;
 XX
 DT 17-FEB-2000 (first entry)
 XX
 DE H. virescens acetyl-choline receptor DNA from clone Hva7-1.
 XX
 KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
 KW neurotransmission; plant protection agent; conductance; AChR; ds.
 XX
 OS Heliothis virescens.

Key Location/Qualifiers
 CDS 335..1825
 /*tag= a
 /product= "acetyl-choline receptor"

DE19819829-Al.
 XX
 PN 11-NOV-1999.
 XX
 PD 04-MAY-1998; 98DE-01019829.
 XX
 PF 04-MAY-1998; 98DE-01019829.
 XX
 PR (FARB) BAYER AG.
 PA
 XX Adamczewski M, Oellers N, Schulte T;
 PI

XX WPI; 2000-014207/02.
 DR P-PSDB; AAY50815.
 XX

PT New nucleic acid encoding a nicotinic acetylcholine receptor from
 XX insects, used to identify potential insecticides.

PS Claim 1a; Page 14-17; 26pp; German.

XX This invention describes a novel nucleic acid (NA) encoding a nicotinic
 CC acetyl-choline receptor (I) from insects which can be used as an
 CC insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
 CC (also vectors containing it, its regulatory regions, and antibodies
 CC directed against (I)-encoded proteins) are used to screen for: (a) plant
 CC protection agents that alter conductance of AChR, potentially useful as
 CC insecticides, or (b) genes which encode polypeptides that are involved in
 CC formation of functionally related AChR in insects. (I) are also used to
 CC isolate and characterize the specified regulatory regions and for
 CC recombinant production of (II). This sequence encodes an acetyl-choline
 CC receptor isolated from *Heliothis virescens*

SQ Sequence 3700 BP; 893 A; 953 C; 944 G; 910 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 9,16e-157 Length: 3700
 Score: 1850.00 Matches: 361
 Percent Similarity: 75.78% Conservative: 30
 Best Local Similarity: 69.96% Mismatches: 49
 Query Match: 45.76% Indels: 76
 DB: 3 Gaps: 7

US-09-303-232-2 (1-770) x AAZ24476 (1-3700)

QY 295 LeuLeuIleTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyrHisGluLysArg 314
 Db 383 CTGCTGTCTGTCTGCTCTGCTGCGCCAGGGGGGCGACGCTCGGGGTACCCAGGAAGCGG 442
 QY 315 LeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeuAsnGluSer 334
 Db 443 CTACTGCACCACTATTGGACCACTACACGTACTGGAGAGGCCCGTCGTCAACGAGAGC 502
 QY 335 AspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleIleAspValAspGluLys 354
 Db 503 GACCCGCTGACGCTCTCTTCGGCCTCAGCTCATCGAGATCATCGACGTCGTCGAGAGAAG 562
 QY 355 AsnGlnLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArg 374
 Db 563 AACGAGCTTTTAAACAAACATCTGGCTAAACTAGATGGATGATATGAATTTGAGG 622
 QY 375 TrpAsnThrSerAspTyrGlyValLysAspLeuArgIleProProHisArgIleTrp 394
 Db 623 TGGAAACATTCAGATTTTCGGCGGGGTCAAAGATTAAAGAGTGCCACCCACAGATATGG 682
 QY 395 LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrThrGlnThr 414
 Db 683 AAACACGACGCTCTTATGTACAAACAGCGCGGACGAAGGGTTCGACAGCATGTCACACG 742
 QY 415 AsnValValValArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSer 434
 Db 743 AACGTGGTGGTGGGAAACAACGCGCTGCTGTCTGTACGTGCGCGCCGCGCATCTTCAAGAGC 802
 QY 435 ThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPhe 454
 Db 803 ACCTGCAAGATCGACATCATCTGTTCCCTTCGACGACCAACGATGCGAGATGAAGTTT 862
 QY 455 GlySerThrThrTyrAspGlyPheGlnLeuAspLeuGlnAspGluThrGlyGly 474
 Db 863 GGCAGCTGGACTTATGATGTTTATAGTTGGATCTCAACTACAGGATGAAGGGGGCGGA 922
 QY 475 AspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLysArg 494
 Db 923 GATATAAGCAGATTTCGACGAATGCGAATGGAGTTAATAGGAGTCCCGCGCAAGCGC 982


```
QY 415 AsnValValArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSer 434
Db 470 AAGTGGTGGTGTGAGAGCGCGGAGTTCCTGTAGTCCACCTGGCATATTCAGAGC 529
QY 435 ThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGluMetLysPhe 454
Db 530 ACATGCAAGATGACATCGCGTGGTTCCCTTCGACGACCAACACTGTGATATGAGTTC 589
QY 455 GlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAspGluThrGlyGly 474
Db 590 GGTAGCTGGACATATGAGCGCAATCAGTTGGATCTGGTCTAAAAGATGAGGACGCGC 649
QY 475 AspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuValProGlyLysArg 494
Db 650 GATCTATCGGACTTCATACAAATGGGAGTGGTATCTAATAGGAATCCAGGCAAAAG 709
QY 495 AsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaIleIle 514
Db 710 AACACATAACATAGCGGTGCTGCCCGAGCCCTAGGTGGAGCTCACCTTCACCATCATG 769
QY 515 IleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysValLeuIleAla 534
Db 770 ATAGAAGACGAACCTTGTACTACTTCTCAACCTGATCGTCCCGTGGTGTGATCTCA 829
QY 535 SerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGly 554
Db 830 TCGATGGCAGCTCCTCGGTTCCACATGCCACGACTCCGGAGAGAACTCACACTTGGG 889
QY 555 ValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAla 574
Db 890 GTCACTATTTCTATCGCTGACGGTGTCTCAACCTGGTAGCGAGACCTGCCACAG 949
QY 575 ThrSerAspAlaValProLeu----- 581
Db 950 GTCTCCGACGCTATCCCCCTGTAGGACGTAFTTCAATTGCATCATGTTTCATGTAAGC 1009
QY 581 ----- 581
Db 1010 TCGTCTGTGTACTGACTGTGTGTACTCAATTACCCACCATCGAACAGCTGATATCAT 1069
QY 582 -----TrpIleArgIleValPheLeuCysTrpLeuProTrpIleLeuArgMet 597
Db 1070 GAAATGCCACAGTGGATAAATCAGTATTCCTACATGGTTGCCATGGATATCGGATG 1129
QY 598 SerArgProGlyArgProLeuLeuLeuGluPheProThrThrProCysSerAspThrSer 617
Db 1130 TCGAGGCCAGGAAGAGATCACCGAAG-----ACTATAATCATGATCAACAGG 1180
QY 618 SerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGluArgSerSerLysSer 637
Db 1181 -----ATGAGGAGCTGGAATCGAATGAGGAGAGGTGCTCGAAGTCC 1219
QY 638 LeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsnCysArgProMetThr 657
Db 1220 TTGTGTCGAATGTTCTAGATATTTGATGATGACTTCAGACACGCGC-----CCTCCGCT 1273
QY 658 ProGly-----GlyThrLeuProHisAsnProAlaPheTyrArgThrVal 672
Db 1274 CCTAACAGTACTCGCTCGACCGGAATTTGGGACCTGGGTGCTCAATATTCGACCGCAT 1333
QY 673 TyrGlyGlnGly-----AspAspGlySerIleGlyProIle 684
Db 1334 TTCGTGCGTGGTTCGTCGTCGTCACATGAGAGAGCTGGGC-----GGCGGCTG 1387
QY 685 GlySerThrArgMetProAspAlaValThrHisHisThrCysIleLysSerThrGlu 704
Db 1388 GGTAGC-----CACCATCGC----- 1402
QY 705 TyrGluLeuGlyLeuIleLeuLysGluIleArgPheIleThrAspGlnLeuArgLysAsp 724
Db 1403 ---GAGTCGACCTACTACTGAGAGAGCTGAGTTCATCATCGGCCAGGATGAAGAGCT 1459
```

```
QY 725 AspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetValValAspArgLeu 744
Db 1460 GATGAGGAAGCGGAGCTGATCAGCGACTGGAAGTTTGTCTGCGATGGTTTGTATAGTTT 1519
QY 745 CysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSerAla 764
Db 1520 TCCCTGTTTCGTGTTTCCACACTTTTCAATCATCGCGACAGTAGCTGTCTGTATCGGCA 1579
QY 765 ProHisIleIleVal 769
Db 1580 CCGCATATCATCGTG 1594
RESULT 4
ABU27131
ID ABL27131 standard; DNA; 984 BP.
XX
AC ABL27131;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 32866.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
PN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE ) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
WPI; 2001-656860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
XX genes from Drosophila and for elucidating cell signaling and cell-cell
XX interactions.
XX
PS Claim 1; SEQ ID NO 32866; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
XX capable of detecting 1000 or more genes from Drosophila. The invention is
XX useful in developmental biology and in elucidating cell signalling and
XX cell-cell interactions in higher eukaryotes for the development of
XX insecticides, therapeutics and pharmaceutical drugs. The invention
XX discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
XX sequences (ABL01840-ABL6175) and the encoded proteins (ABBS7737-
XX ABH72072). The sequence data for this patent did not form part of the
XX printed specification, but was obtained in electronic format directly
XX from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 984 BP; 332 A; 268 C; 225 G; 159 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 8.57e-126 Length: 984
Score: 1498.50 Matches: 297
Percent Similarity: 98.69% Conservative: 4
Best Local Similarity: 97.38% Mismatches: 3
Query Match: 37.06% Indels: 1
DB: 4 Gaps: 1
US-09-303-232-2 (1-770) x ABL27131 (1-984)
QY 1 MetLysAsnAlaGlnLeuLysLeuThrGluValAspAspGluLeuTrpLeuAlaVal 20
```

Db 1 ATGAATAATGCACAACTGAACTGACTGAACTTGCACGATGATGAGCTGTGGCTGGCAGTA 60
 Qy 21 ArgLeuAlaHisCysSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 40
 Db 61 AGATTAGCGCACTGACGAGCAACATCAGCAGCAGTAGCAGCAACAGAACCCAGCAGC 120
 Qy 41 AsnGlnArgHisAsnGlnGlnLeuThrThrLeuGlnProArgSerLeuSerThrLysHis 60
 Db 121 AACAGAGGCAACACCGCAACTCACAACACTGCACCAAGAGAGCTTAAGTACAACAC 180
 Qy 61 HisSerAsnLeuAlaSerGluGlnHisAsnSerGlnGlnGlnGlnProAlaSerLysAsp 80
 Db 181 CACAGCAACATTCGACGAGCAGCAGCAATAGCCAGCAACAGGAGCCAGCATCGAAGGAC 240
 Qy 81 GluAspValAlaAsnHisGlyArgSerAsnAspGlnGlnThrHisLeuGlnGlnLeuAsp 100
 Db 241 GAGGATGTAGCAACACCGTAGAGCAATCAGCAGCAGCAGCAGCATCTGCAACAGCTAGAC 300
 Qy 101 SerSerAsnMetLeuSerProLysThrAlaAlaAlaAlaThrAlaAlaGlyAspGluAla 120
 Db 301 AGCAGCAACATGTTGTGCCAAAGACAGCCGAGCAGCAACTGCTGCCGCGGATGAAGCA 360
 Qy 121 ThrThrGlnGlnProThrAsnLeuArgLeuCysAlaArgLysArgGlnArgLeuArgArg 140
 Db 361 ACAACCAACCAACCAACCAATAGACTGTGTGCAGCAAGCGCAACAGATTGGCTGCGC 420
 Qy 141 ArgArgLysArgLysProAlaThrProAsnGlnThrAspLysLysGlnGlnGlnLeu 160
 Db 421 CGACGAAAGAAACACCGCAACCCCAACCAAGCAATATCAAGAAACACAGCAACTT 480
 Qy 161 SerMetProProPheLysThrArgLysSerThrAspThrThrSerThrProAlaAlaThr 180
 Db 481 AGCATGCTCTCTTCAAACGGCGCAATCCACGGACACTAGCAGCACACAGCAATA 540
 Qy 181 ThrSerCysProThrAlaThrTyrMetGlnCysArgAlaSerAspAsnGluPheSerLys 200
 Db 541 ACCAGCTGTCCGACAGCCACTACATGCAATGTGCAGCGCAGCAACATGAGTTTCAGTAT 600
 Qy 201 ProLysSerArgHisAspArgValSerThrAlaThrPheAlaTrpValLeuHisValLeu 220
 Db 601 CCGATATCGAGACATGATAGATATCCACGGCCACATTCGCTGGGTGTGATGTGCTG 660
 Qy 221 GlnValLeuLeuValSerLeuGlnGlnTrpGlnLeuHisValGlnGlnArgSerValLeu 240
 Db 661 CAGGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 720
 Qy 241 LeuPheArgArgLeuAlaAlaSerThrLeuAlaPheLysSerThrLeuGlySerPheAla 260
 Db 721 CTGTTTCAGAAAGATCGCAGCGAGCACCATCGCTTCTTCTTCTTCTTCTTCTTCTTCT 780
 Qy 261 AlaGlnLeuLysAsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSer 279
 Db 781 GCGCAACTCAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 840
 Qy 280 SerThrGlnLeuLeuAsnGlyLeuAsnLysHisSerThrPheLeuLeuLeuLeuLeu 299
 Db 841 AGCAGCAATATTAACCGACTTAATAAACAACACTCATGCTATATTTTATTGATATATTG 900
 Qy 300 AsnLeuSerAlaLys 304
 Db 901 AATTTATCTGCTAAA 915

RESULT 5

ABL07231

ID ABL07231 standard; cDNA; 1540 BP.

XX AC

XX ABL07231;

XX 26-MAR-2002 (first entry)

DT XX

DE Drosophila melanogaster expressed polynucleotide SEQ ID NO 16175.

XX XX

KW Drosophila; developmental biology; cell signalling; insecticide;

pharmaceutical; gene; ss.
 Drosophila melanogaster.
 WO200171042-A2.
 27-SEP-2001.
 23-MAR-2001; 2001WO-US009231.
 23-MAR-2000; 2000US-0191637P.
 11-JUL-2000; 2000US-00614150.
 (PEKE) PE CORP NY.
 Venter JC, Adams M, Li PWD, Myers EW;
 WPI; 2001-656860/75.
 P-PSDB; ABB63128.

New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.

Claim 1; SEQ ID NO 16175; 2lpp + Sequence Listing; English.

The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (AB116176-AB130511), expressed DNA sequences (AB101840-AB16175) and the encoded proteins (ABB57737-ABB72072). The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences

Sequence 1540 BP; 412 A; 390 C; 376 G; 362 T; 0 U; 0 Other;

Alignment Scores:

Pred. No.: 1,33e-122 Length: 1540
 Score: 1466.50 Matches: 302
 Percent Similarity: 65.09% Conservative: 41
 Best Local Similarity: 57.31% Mismatches: 73
 Query Match: 36.27% Indels: 112
 DB: 4 Gaps: 8

US-09-303-232-2 (1-770) x ABL07231 (1-1540)

Qy 291 SerTrpPheLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeuLeu 310
 Db 71 TCGCTGTTGTCTCTTCTGATCTTTCTGGCGATAATTAAGAAAGCTGT---CAAGGACCT 127
 Qy 311 HisGluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProVal 330
 Db 128 CATGAAAGCGCTGCTGAACCATCTGCTGTGCCACTACAAATCGCTGAGGAGCACCCTG 187
 Qy 331 LeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnLeuLeuAsp 350
 Db 188 GCCATGATCGAGGCCCTCGAGGTTAAGTTCCGACTGACGCTGCGAGCAGATCATCGAC 247
 Qy 351 ValAspGluLysAsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAsp 370
 Db 248 GTGGATGAAAGAAATCAGATTCTGACCAAAATGCGTGGTTAAATTTGAGGTGGAACGAC 307
 Qy 371 MetAsnLeuArgTrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgLysProPro 390
 Db 308 TACAATCTCGCTGGATGAATGAACGGAATACGGGGGTCAAGGATCTACGAATCACGCC 367
 Qy 391 HisArgLysTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGly 410
 Db 368 AACACGCTGTGAAGCCCGACGCTGCTCATGTACACAGCGCGGATGAGGGATTCTGATGGC 427

QY 411 ThrTyrGlnThrAsnValValValArgAsnAsnGlySerCysLeuTyrValProGly 430
Db 428 ACGTATCACCAACATTGTGTCTCAACATAACGCGAGTTGTCTGTACGTGCCCCCTGGT 487
QY 431 IlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCys 450
Db 488 ATCTTTCAAGAGACATGCAAGATAGACATCACGTGTTCCCATTTGATGACCAACATTGC 547
QY 451 GluMetLysPheGlySerThrTrpTyrAspGlyPheGlnLeuAspLeuGlnAsp 470
Db 548 GAAATGAATTCGGTAGTTGGACTTACGATGAAATCAGTTGGATTTGGTTGAATTCC 607
QY 471 GluThrGlyGlyAspIleSerThrValLeuAsnGlyGluTrpGluLeuGly--- 489
Db 608 GAAGATGAGGGGATCTTCCGATTTCATAACAAATGGCGAGTGTACTTGTCTGGTTAC 667
QY 490 -ValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIle 509
Db 668 CATGCCGGGAAGAAGATAGATAGTCTACGCTCTGCCAGAACCATATGTCGATAT 727
QY 509 eThrPheAlaIleIleIleArgArgThrLeuTyrTyrPhePheAsnLeuIleIlePr 529
Db 728 CACCTTACTATACAAATCTGCGCGTACATATATTATTTTCAATTTAATTTGTCGC 787
QY 529 oCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyG1 549
Db 788 ATGTGTGCTAATCTCATGATGGCCCTACTGGCTTTCATATGCGCGGATTCGGCGCA 847
QY 549 uLysLeuSerLeuGlyValThrIleLeuSerLeuThrValPheLeuAsnMetValAl 569
Db 848 GAAACTGACGTGGCGTAACTACTACTATCATTAACAGTATTTCTAAACCTTGTGCG 907
QY 569 aGluThrMetProAlaThrSerAspAlaValProLeu----- 581
Db 908 CGAGTCCATGCCGACGACGTCGGATGCTGTTCTCTCT-TATAGGTACACACAGATATT 966
QY 581 ----- 581
Db 967 AAAACGCTATAGCTTAAAGTACGACGATCGGTGGCTATCGACCTACTACCACTG 1026
QY 581 ----- 581
Db 1027 TCGAGCAGAAAGCACCTACTTCAATTGCATCATGTTCATGTGCGCTCGTGGTGTCT 1086
QY 582 -----Tr 582
Db 1087 GACAGTAGTGTGCTCACTACCATCCACACGCGGACATTCACGAGATGCCACCGTG 1146
QY 582 pIleArgIleValPheLeuCysTrpLeuProTrpIleLeuArgMetSerArgProGlyVar 602
Db 1147 GATCAAGTCGGTTTCTCTACAAATGGCTGCCCTGGATCTTGGATGGTGCACCGGTG 1206
QY 602 qProLeuIleLeuGluPheProThrThrProCysSerAspThrSerSerGluArgLysH1 622
Db 1207 CAGATT-----ACACGCAAAAC 1224
QY 622 sGlnIleLeuSer-----AspValGluLeuLysGluArgSerSerLysSerLe 638
Db 1225 AATACTATTAAAGCAATCGCATGAAGAGCTGGAGCTAAAGAGCGCTCCCTCCAAATCCCT 1284
QY 638 uLeuAlaAsnValLeuAspIleAspAspPheArgHisAsnCysArgProMetThrPr 658
Db 1285 GCTGGCCAAATGTCCTGCATCGACATCGACGACGATTCGCGCACACAATA----- 1330
QY 658 oGlyGlyThrLeuProHisAsnProAlaPheTyrArgThrValTyrGlyGlnGlyAspAs 678
Db 1331 -----TCTGGCTCCCA 1341
QY 678 pGlySerIleGlyProIleGlySerThrArgMetProAspAlaVal-----ThrHisIsth 697
Db 1342 AACCGCATTTGGCTCGCGCCAGCTTCGGTCGGCCCAACACGCTGGAGGAGATCACAC 1401
QY 697 rCysIleLysSerSerThrGluTyrGluLeuGlyLeuIleLeuLysGluIleArgPheI1 717

Db 1402 GCCATC---GGCTCAATCAAAAGATCTTCATTAATCTCAAGAATTGCAATTAT 1458
QY 717 eThrAspGlnLeuArgLysAspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAl 737
Db 1459 TACGGCGGGATGCGCAAGAGCTGACGACGAGCGAATGATGCGGATTTGAAGTTCG 1518
QY 737 aAlaMetValValAspArg 743
Db 1519 GGCAATGTTGTGGATAGG 1537
RESULT 6
ABL27130
ID ABL27130 standard; DNA; 3144 BP.
XX
AC ABL27130;
XX
DT 26-MAR-2002 (first entry)
XX
DE Drosophila melanogaster genomic polynucleotide SEQ ID NO 32863.
XX
KW Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ds.
XX
OS Drosophila melanogaster.
XX
FN WO200171042-A2.
XX
PD 27-SEP-2001.
XX
PF 23-MAR-2001; 2001WO-US009231.
XX
PR 23-MAR-2000; 2000US-0191637P.
XX
PR 11-JUL-2000; 2000US-00614150.
XX
PA (PEKE) PE CORP NY.
XX
PI Venter JC, Adams M, Li PWD, Myers EW;
XX
DR WPI; 2001-658860/75.
XX
PT New isolated nucleic acid detection reagent for detecting 1000 or more
PT genes from Drosophila and for elucidating cell signaling and cell-cell
PT interactions.
XX
PS Claim 1; SEQ ID NO 32863; 21pp + Sequence Listing; English.
XX
CC The invention relates to an isolated nucleic acid detection reagent
CC capable of detecting 1000 or more genes from Drosophila. The invention is
CC useful in developmental biology and in elucidating cell signalling and
CC cell-cell interactions in higher eukaryotes for the development of
CC insecticides, therapeutics and pharmaceutical drugs. The invention
CC discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
CC sequences (ABL01840-ABL16175) and the encoded proteins (AB57737-
CC AB572072). The sequence data for this patent did not form part of the
CC printed specification, but was obtained in electronic format directly
CC from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 3144 BP; 1056 A; 676 C; 680 G; 732 T; 0 U; 0 Other;
Alignment Scores:
Pred. No.: 3,12e-121 Length: 3144
Score: 1456.50 Matches: 300
Percent Similarity: 86.89% Conservative: 5
Best Local Similarity: 85.47% Mismatches: 12
Query Match: 36.03% Indels: 35
DB: 4 Gaps: 3
US-09-303-232-2 (1-770) x ABL27130 (1-3144)
QY 1 MetLysAsnAlaGlnLeuLysLeuThrGluValAspAspGluLeuTrpLeuAlaVal 20
Db 1001 ATGAAATATGCAACTGAATGACTGAAGTTGACGATGATGAGCTGGCTGGCAGTA 1060

```
QY 21 ArgLeuAlaHisCysSerSerAsnPheserSerSerSerThrArgThrThrSerSer 40
Db 1061 AGATTAGCGCACTCAGCAGCAACATCAGCAGCAGTAGCAGCAACAGAACCCACGAGCAGC 1120
QY 41 AsnGlnArgHisAsnGlnGlnLeuThrThrLeuGlnProArgSerLeuSerThrLysHis 60
Db 1121 AACAGAGGCCACACCGCACTCACAACACTGCAACCAAGAGCTTAAGTACAAACAC 1180
QY 61 HisSerAsnIleAlaSerGlnGlnHisAsnSerGlnGlnGlnProAlaSerLysAsp 80
Db 1181 CACAGCAACATTGCAAGCAGCAGCAGCACAATAGCCAGCAACAGGAGCCAGCATCGAAGCAGC 1240
QY 81 GluAspValAlaAsnHisGlyArgSerAsnAspGlnGlnThrHisLeuGlnGlnLeuAsp 100
Db 1241 GAGGATGTAGCCACACCGGTAGAGCAATGACCCAGCAGCAGCATCTGCAACAGCTAGC 1300
QY 101 SerSerAsnMetLeuSerProLysThrAlaAlaAlaAlaThrAlaAlaGlyAspGluAla 120
Db 1301 AGCAGCAACATGTTGTGCCAAAGACAGCAGCAGCAACTGTGCGCGCGATGAAGCA 1360
QY 121 ThrThrGlnGlnProThrAsnIleArgLeuCysAlaArgLysArgGlnArgLeuArgArg 140
Db 1361 ACAACCCCAACCAACCAACATAGACTGTGTGCACGCAAGCGCAACAGATTGCGTCGC 1420
QY 141 ArgArgLysArgLysProAlaThrProAsnGluThrAspIleLys----- 155
Db 1421 CGACGAAAGAAACCAACGACCAACCCCAACGAAACAGATATCAA--GGTAACATTAATAA 1479
QY 155 ----- 155
Db 1480 CATCATATTAACTAAATATAGAAATTTAGAAAAATTAATTCACCCCTCAGCACTTGAAT 1539
QY 156 -----LysGlnGlnGlnLeuSerMetProPhenylsThrArg 168
Db 1540 TTGGTCTTTCTTACAATTCAGAAACACAGCAACTTAGCATGCTCCCTTCAAAACGGCGC 1599
QY 169 LysSerThrAspThrSerThrProAlaAlaThrThrSerCysProThrAlaThrTyr 188
Db 1600 AAATCCAGGACACTCAGCAGCAGCAGCAATAACCACTGTCCGACACCCACTTAC 1659
QY 189 MetGlnCysArgAlaSerAspAsnGluPheSerIleProIleSerArgHisAspArgVal 208
Db 1660 ATGCAATGTCGAGCGAGCAGCAATGAGTTTCAGTATTCGATATCGAGACATGATAGAT 1719
QY 209 SerThrAlaThrPheAlaTrpValLeuHisValLeuGlnValLeuValSerLeuGln 228
Db 1720 TCCACGGCCACATTCGCTGGGTGTGCAATGTGCTGAGGTGCTGCTGCTGCTGCA 1779
QY 229 GlnTrpGlnLeuHisValGlnGlnArgSerValLeuLeuPheArgArgIleAlaAlaSer 248
Db 1780 CAGTGGCAACTTCAGTGCAACAGCAGCATCGTGCTACTGTTTCAAGAGGATCGCAGGAGC 1839
QY 249 ThrIleAlaPheIleSerTyrLeuGlySerPheAlaAlaGlnLeuLysAsnSerSerSer 268
Db 1840 ACCATCGCCTTCATTTCTTATTTAGCAGCTTTGACGCGCAACTGAGGAGCAGCAGCAT 1899
QY 269 SerSerSer---SerSerAsnSerSerAsnSerSerThrGlnIleLeuAsnGlyLeu 287
Db 1900 AGCAGCAGCAACAGCAGCAACAGCGGCAACAGCAGCAGCAGCAATATTAACCGGACTT 1959
QY 288 AsnLysHisSerTrpIlePheLeuLeuIleTyrLeuAsnLeuSerAlaLys---ValCys 306
Db 1960 AATAACACTCATGATATTTTATTTGATATATTTGATATTTATCTCTAAAGGTAGTGC 2019
QY 307 LeuAlaGlyTyrHisGlnLysArgLeuLeuHis 317
Db 2020 CATTTAAATATCTCGAGTGTGATGCAACAT 2052
Db 2020 CATTTAAATATCTCGAGTGTGATGCAACAT 2052
RESULT 7
ABLI13733
ID ABLI13733 standard; cDNA; 936 BP.
XX
```



```

Qy 572 MetProAlaThrSerAspAlaValProLeuTrpIleArgIleValPheLeuCysTrpLeu 591
Db |||||||TCTTCTCGATGCGTGCATG-----624
Qy 592 ProTrpIleLeuArgMetSerArgProGlyArgProLeuIleLeuGluPheProThrThr 611
Db -----624
Qy 612 ProCysSerAspThrSerSerGluArgGlyHisGlnIleLeuSerAspValGluLeuLys 631
Db |||||CTGGAACCTGACATAATCAATTTAATCAAAATCAAA-----660
Qy 632 GluArgSerSerLysSerLeuAlaAsnValLeuAspIleAspAspPheArgHis 651
Db -----660
Qy 652 AsnCysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyrArgThr 671
Db -----660
Qy 672 ValTyrGlyGlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAsp 691
Db -----660
Qy 692 AlaValThrHisHisThrCysIleLysSerSerThrGluTyrGluLeuGlyLeuIleLeu 711
Db -----660
Qy 712 LysGluIleArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsnAspIleAla 731
Db -----CAGCTACGTAAAGATGACGAGTGCAATGACATTGCC 696
Qy 732 AsnAspTrpLysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrMet 751
Db |||||||AATGATTGGAATTTCAGCTATGCTGTGACAGCTGTGCTTATCATATTCACAATG 756
Qy 752 PheAlaIleLeuAlaThrIleAlaValLeuLeuSerAlaProHis 766
Db |||||||TTCACAATATTAGCCACATAGCTGTACTATCAGACCCACAT 801

RESULT 9
ID AAT59196 standard; cDNA; 2769 BP.
AC AAT59196;
XX
DT 17-JUN-1997 (first entry)
DE Neuronal alpha-bungarotoxin binding protein alpha subunit cDNA.
KW Neuronal alpha-bungarotoxin binding protein alpha 1; cholinergic;
KW ligand binding; ion channel; ss.
XX
OS Gallus sp.
XX
FH Key Location/Qualifiers
FT CDS 71..1513
FT /tag= a
FT sig_peptide 71..136
FT /tag= b
FT mat_peptide 137..1510
FT /tag= c
XX
PN US5599709-A.
XX
PD 04-FEB-1997.
XX
PF 28-SEP-1989; 89US-00413947.
XX
PR 28-SEP-1989; 89US-00413947.
XX
PA (SALK ) SALK INST BIOLOGICAL STUDIES.

```

```

XX Lindstrom JM, Schoepfer RD;
XX WPI; 1997-118297/11.
XX P-PSDB; AAM12368.
XX New isolated neuronal alpha-bungarotoxin-binding protein DNA - used to
XX screen cholinergic agents and other drugs which may affect ligand
XX binding, ion channel or other activities of the protein.
XX Claim 1; Fig 2A-B; 18pp; English.
XX
CC 2 cDNA clones (AAT59196 and AAT59197) respectively code for the alpha
CC subunit (AAM12368) and alpha2 subunit (AAM12369) of chick neuronal alpha-
CC bungarotoxin binding protein (ABBP). They were isolated from an 18-day
CC embryo chick brain cDNA library using a probe (see also AAT59198) based
CC on the N-terminal amino acid sequence of chicken brain ABBP. The probe
CC isolated partial clone pCh29-1, which encoded the N-terminal portion of
CC alpha1. A subclone, pCh29-3 (ATCC 40641), was used to rescreen the
CC library, yielding clone pCh31-1 (ATCC 40640), which encoded the entire
CC alpha2 sequence. A probe based on the C-terminal region of pCh31-1 was
CC used to obtain clone pCh34-1 (ATCC 40639), encoding the C-terminal portion
CC of alpha1. The cDNA clones can be used as probes to identify further ABBP
CC subunits, and in the recombinant prodn. of ABBP
XX
SQ Sequence 2769 BP; 790 A; 604 C; 571 G; 804 T; 0 U; 0 Other;

Alignment Scores:
Pred. No.: 3.18e-96 Length: 2769
Score: 1179.50 Matches: 237
Percent Similarity: 61.08% Conservative: 80
Best Local Similarity: 45.66% Mismatches: 127
Query Match: 29.17% Indels: 75
DB: 2 Gaps: 9

US-09-303-232-2 (1-770) x AAT59196 (1-2769)
Qy 295 LeuLeuIleTyrLeuAsnLeuSerAla-----LysValCysLeuAlaGlyTyrHis 311
Db |||||||TCTGCTGCTGCGCGCGCGGCGTCTGCGAGTCCCTGCAAGGAGAGTTC 145
Qy 312 GluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeu 331
Db |||||||CAAAGGAAGCTGTACAGGAGCTGTGAAGAACTACAACTCTGGAACGACGAGTGCA 205
Qy 332 AsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAspVal 351
Db |||||||AATGACTCCAGCGCTCAGTGTCTATTCTCTCAGCCTCATCGAGATCATGATGTG 265
Qy 352 AspGluLysAsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMet 371
Db |||||||GATGAAAAGAAATCAAGTATTAAACAACAACATCTGGCTACAAATGTCTGGACAGATCAT 325
Qy 372 AsnLeuArgTrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIleProHis 391
Db |||||||TACTTACAGTGAATGTCTGCAATACCTCGAGTGAAGAGCGTCCGTTTCTCTGATGA 385
Qy 392 ArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThr 411
Db |||||||CTGATTGGAGCCAGATATTCTTCTCTATACATGCTGATGAAGATTGTGATCTACA 445
Qy 412 TyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProGlyIle 431
Db |||||||TTTCACACTAATGTTTGTAGTCAATTTCTTGGGACACTGCCAATATCTCCACGACGATA 505
Qy 432 PheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGluArgCysGlu 451
Db |||||||TTTAAAGCTCATGCTACATAGCGTGGTGGTTTCCATTGATGTTTCAGAAAGTCAAT 565
Qy 452 MetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAspGlu 471
Db |||||||CTGAAGTTTGGATCTTGGACATATGAGGCTGGTCTCTAGACTTACAAATGCAAGAA--- 622

```

QY 472 ThrGlyGlyAspIleSerSerTyrValIleuAsnGlyGluThrGluLeuGlyValPro 491
 Db 623 -----GCAGATATCCGGCTATATTCAATATGGAGTGGGATTTAGTAGAATTCCT 676
 QY 492 GlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPhe 511
 Db 677 GGAAGAGAACTGAGAGCTTTTATGAGTGTGTAAAGAACCATACCCAGATATCACATTC 736
 QY 512 AlaIleIleIleArgArgGluThrLeuTyrTyrPheAsnLeuIleIleProCysVal 531
 Db 737 ACAGTAACCATGAGACGAGCAACTCTCTACTACGGCTCAACCTTCTTATTCCTGTGTA 796
 QY 532 LeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeu 551
 Db 797 CTGATATCAGACACTGCTTATTAGTCTTCTGCTTCCAGCAGACTCAGGAGAAAGATC 856
 QY 552 SerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThr 571
 Db 857 TCACGTAGGTATAACAGATTTTATTCTCTCACCCTCTTCATGTTACTCGTGGCTGAAT 916
 QY 572 MetProAlaThrSerAspAlaValProLeu----- 581
 Db 917 ATGCCAGCAATCTGATTCGTGCTTAAATGCTCAGTATTTTGGCAGCACCATGATT 976
 QY 581 ----- 581
 Db 977 ATGTGGCTCTCTGTGTTGTGCATGTTATCGTCTCAATACATCATCAGATCCA 1036
 QY 582 -----TrpIleArgIleValPheLeuCysTyrLeuProTyrIle 594
 Db 1037 GATGGGGGAAAAATGCCATAATGGACAAGATCATCTCTGAATTCGTGTGCTGTT 1096
 QY 595 LeuArgMetSerArgProGlyArgProLeuIleLeuGluPheProThrThrProCysSer 614
 Db 1097 CTGAGGTGTAAGACACAGG-----GAAGATAAAGTCGTCCGCTGT 1141
 QY 615 AspThrSerSerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGluArgSer 634
 Db 1142 CAACATAAACAGCGCCGATGACG-----CTGTACAGCATGGAGTGAACACTGTGAGT 1195
 QY 635 SerLysSerLeuLeu-----AlaAsnValLeuAspIleAspAspPheArgHis----- 651
 Db 1196 GGTGAGCAATGCAATGTAATGGAAACATGCTGTATATT-----GGGTTTCGAGGCTGGAT 1249
 QY 652 -----AsnCysArgProMetThrProGlyThrLeuProHisAsnProAlaPheTyr 669
 Db 1250 GGGGTTCACTGCACACCCACCTGATTCAGGGGTGATC----- 1288
 QY 670 ArgThrValTyrGlyGlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMet 689
 Db 1289 -----TGTGGAGGATGACCTGTTACCACACA 1315
 QY 690 ProAspAlaValThrHisHisThrCysIleLysSerSerThrGluTyrGluLeuGlyLeu 709
 Db 1316 GAGGAGAAATCTTCGACAGTGGCCACCCCTCTGAGGCGCAGCCAGATTTGGCTAAG 1375
 QY 710 IleLeuLysGluIleArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsnAsp 729
 Db 1376 ATCTGGAAGAGTGCATGATCATTCGAAAGGTTTCAGACCCAGGATGAAGAAAGACC 1435
 QY 730 IleAlaAsnAspTrpLysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePhe 749
 Db 1436 ATTGCAACGAATGGAAGTTTGAGCGCTCTGTAGTAGATCGGCTGCTGTGATGGCATTT 1495
 QY 750 ThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSerAlaProHisIleIle 768
 Db 1496 TCGGCTCTTCACCATCATTTTGTAAATTTGGCATCTTAATGTGAGCAGCAACAACTTTGTA 1552
 RESULT 10
 ID AAT59197
 XX
 AC AAT59197;

XX 17-JUN-1997 (first entry)
 DT
 XX
 DE Neuronal alpha-bungarotoxin binding protein alpha2 subunit cDNA.
 XX
 XX Neuronal alpha-bungarotoxin binding protein alpha 2; cholinergic;
 KW ligand binding; ion channel; ss.
 XX
 OS Gallus sp.
 XX
 FH Key Location/Qualifiers
 CDS 56..1501
 FT sig_peptide /*tag= a
 FT 56..145
 FT mat_peptide /*tag= b
 FT 146..1498
 FT /*tag= c
 PN US5599709-A.
 XX
 PD 04-FEB-1997.
 XX
 PF 28-SEP-1989; 89US-00413947.
 XX
 PR 28-SEP-1989; 89US-00413947.
 XX
 XX (SALK) SALK INST BIOLOGICAL STUDIES.
 PA Lindstrom JM, Schoepfer RD;
 PI WPI; 1997-118297/11.
 XX P-PSDB; AAW12369.
 DR
 DR New isolated neuronal alpha-bungaro-toxin-binding protein DNA - used to
 XX screen cholinergic agents and other drugs which may affect ligand
 XX binding, ion channel or other activities of the protein.
 PS Claim 1; Fig 3A-B; 18pp; English.
 XX
 CC 2 cDNA clones (AAT59196 and AAT59197) respectively code for the alpha1
 CC subunit (AAW12368) and alpha2 subunit (AAW12369) of chick neuronal alpha-
 CC bungarotoxin binding protein (ABBP). They were isolated from an 18-day
 CC embryo chick brain cDNA library using a probe (see also AAT59198) based
 CC on the N-terminal amino acid sequence of chicken brain ABBP. The probe
 CC isolated partial clone pCh29-1, which encoded the N-terminal portion of
 CC alpha1. A subclone, pCh29-3 (ATCC 40641), was used to rescreen the
 CC library, yielding clone pCh31-1 (ATCC 40640), which encoded the entire
 CC alpha2 sequence. A probe based on the C-terminal region of pCh31-1 was
 CC used to obtain clone pCh34-1 (ATCC 40639), encoding the C-terminal portion
 CC of alpha1. The cDNA clones can be used as probes to identify further ABBP
 CC subunits, and in the recombinant prodn. of ABBP
 XX
 SQ Sequence 2101 BP; 582 A; 406 C; 454 G; 659 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 1,36e-95 Length: 2101
 Score: 1170.50 Matches: 244
 Percent Similarity: 58.02% Conservatives: 85
 Best Local Similarity: 43.03% Mismatches: 129
 Query Match: 28.95% Indels: 109
 DB: 2 Gaps: 13
 US-09-303-232-2 (1-770) x AAT59197 (1-2101)
 QY 265 AsnSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerSerThrGlnIleLeu 284
 Db 2 AATCATCTCTTTTACCTCTGCTCAGAAAGGATAGACAGCCAGTTCGACAAATGCTT 61
 QY 285 -----AsnGlyLeuAsnLysHisSerTrpIlePheLeu 295
 Db 62 ACTGAAAAGTGTGGGATTTTTTTTATTCTGGACTGTGCTGGGCACTCTGTTCTG 121
 QY 296 LeuIleTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyrHisGluLysArgLeu 315

Db	1157	----	AGTCATCCCAACACCACTCCACGCTGAAAAACACAGAGATG	-----	1199
Qy	636	LySerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsnCysArgPro	655		
Db	1199	-----	AATGTCCTA	-----	1207
Qy	656	MetThrProGlyGlyThrLeuProHisAsnProAlaPheTyrArgThrValTyrGly	674		
Db	1208	-----	CCTGGG	-----	1249
Qy	675	-----	GlnGlyAspAsp	-----	683
Db	1250	CACACAATGAAATCAATGCTGCCCCAGAACAAATGATCTGGGCAGCAAGAGTGAAG	1309		
Qy	684	IleGlySerThrArgMetProAspAlaValThrHisHisThrCysIleLysSer	701		
Db	1310	ATTACTTGCCCTTGTCAGAGAT	-----	AACGAGCATGTTCAAAAAAAGCTTTAATG	1363
Qy	702	SerThrGluTyrGluLeuGlyLeuIleLeuLysGluIleArgPheIleThrAspGlnLeu	721		
Db	1364	GATACCATCCAGTGTGTGAAGATACTGGAGGAAGTTTCAGTTCATAGCAATGCGGCTC	1423		
Qy	722	ArgLysAspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaIleMetValVal	741		
Db	1424	AGAAGCAAGATGAGGCTGAAGAGATCTGCAGTGAGTGAAGAGTTTCAGCTGCTGTCATA	1483		
Qy	742	AspArgLeuCysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeu	761		
Db	1484	GACAGATTATGCTGGTGGTTTACCTTTTGGCCATCATTTTGCACATTTTACAATACTC	1543		
Qy	762	LeuSerAlaProHisIleIle	768		
Db	1544	ATGTCGTCTCCCAACTTTATA	1564		
RESULT 11					
AA	C58395	ID AAC58395 standard; cDNA; 1509 BP.			
AC	AC	AAC58395;			
AC	XX	29-JAN-2001 (first entry)			
DT	XX	Human PRO2145 nucleotide sequence SEQ ID NO:76.			
DE	XX	Human; tumour; diagnosis; neoplastic disease; neoplastic cell growth;			
KW	XX	proliferation; tumorigenesis; identification; cancer; cytostatic;			
KW	XX	neotropic; neuroprotective; antiinflammatory; immunosuppressive;			
KW	XX	immunostimulant; antiangiogenic; leukaemia; lymphoid malignancy;			
KW	XX	neural disorder; glial disorder; astrocytal disorder; angiogenic;			
KW	XX	hypothalamic disorder; glandular disorder; macrophagal disorder;			
KW	XX	epithelial disorder; stromal disorder; blastocoelec disorder;			
KW	XX	inflammatory disorder; immunologic disorder; ss.			
OS	XX	Homo sapiens.			
XX	XX	WO2000053755-A2.			
XX	XX	14-SEP-2000.			
XX	XX	06-JAN-2000; 2000WO-US0000376.			
PF	XX	08-MAR-1999; 99WO-US005028.			
PR	XX	02-JUN-1999; 99WO-US012252.			
PR	XX	23-JUN-1999; 99US-0141037P.			
PR	XX	07-JUL-1999; 99US-0143048P.			
PR	XX	26-JUL-1999; 99US-0145698P.			
PR	XX	30-NOV-1999; 99WO-US028113.			
PR	XX	20-DEC-1999; 99WO-US030911.			
PR	XX	05-JAN-2000; 2000WO-US0000219.			
PA	XX	(GETH) GENENTECH INC.			
PI	XX	Ashkenazi AJ, Baker KP, Goddard A, Gurney AL, Hillan KJ Roy MA;			


```
Db 1177 GACGGCGTGCACGTGTCCCG---ACCCCC-----1203
QY 669 TyrArgThrValTyrGlyGlnGlyAspGlySerIleGlyProIleGlySerThrArg 688
Db 1204 -----GACTCTGGGCTAGTG-----TGTGGCGCG 1227
QY 689 MetProAspAlaValThrHis-----HisThrCysIleLysSerSerThr 703
Db 1228 ATGGCTGCTCCCGACGACGATGACACCTCTGTCAGCGGGGCAACCCCGAGGGG 1287
QY 704 GlutyrGluLeuGlyLeuLeuLysGluLeuArgPheIleThrAspGlnLeuArgLys 723
Db 1288 GACCCGGACTTGGCCAGATCTCTGGAGAGTCCGCTACATTGCCAATCGCTCCGCTGC 1347
QY 724 AspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetValValAspArg 743
Db 1348 CAGGACGAAAGCAGGCGGTCTGCAGCGAGTGAAGTTGCCCGCTGTGTGGTGCAGCG 1407
QY 744 LeuCysIleIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeuSer 763
Db 1408 CTGTGCTCATGGCCTTCTGGGTCTTCCATCATCTGCACCATCGGCATCTGTATGTCG 1467
QY 764 AlaProHisIleIle 768
Db 1468 GCTCCCAACTTCGTG 1482

RESULT 13
AAV12197
ID AAV12197 standard; cDNA; 1876 BP.
XX
AC AAV12197;
XX
XX
DT 14-MAY-1998 (first entry)
XX
DE Human neuronal nicotinic acetylcholine receptor alpha-7 subunit cDNA.
KW Human; neuronal nicotinic acetylcholine receptor; alpha-7 subunit;
KW brain tissue; screening; NACHR; antibody; ds.
XX
OS Homo sapiens.
XX
FH Key
FH 73.1581
FH CDS
FT /tag= a
FT /product= "neuronal nicotinic acetylcholine receptor
FT alpha-7 subunit"
XX
PN W09420617-A2.
XX
PD 15-SEP-1994.
XX
PF 08-MAR-1994; 94MO-US002447.
XX
PR 08-MAR-1993; 93US-00028031.
XX
PA (SALK ) SALK INST BIOTECHNOLOGY IND ASSOC.
XX
PI Elliott KJ, Ellis SB, Harpold MM;
DR WPI; 1994-303024/37.
DR P-PSDB; AAW44153.
XX
PT Human neuronal nicotinic acetylcholine receptor subunits and DNA - also
PT transformed cells useful for screening cpds. which modulate activity of
PT the receptor.
XX
PS Claim 8; Page 78-79; 99pp; English.
XX
CC The present sequence encodes a human neuronal nicotinic acetylcholine
CC receptor (NACHR) subunit. The cells expressing the alpha and/or beta
CC NACHR subunits may be used in a method of screening compounds to identify
CC any which modulate the activity of human neuronal NACHR. Subunit specific
CC antibodies may be used to monitor the distribution and expression density
```

```
CC of various subunits in normal vs diseased brain tissues. Testing of
CC single receptor subunits or specific receptor subunit combinations with a
CC variety of potential agonists or antagonists provides information with
CC respect to the function and activity of the individual subunits and
CC should lead to the identification and design of compounds that are
CC capable of very specific interaction with one or more receptor subtypes.
CC The resulting drugs should exhibit fewer unwanted side effects than drugs
CC identified e.g. screening with cells that express a variety of subtypes
XX
SQ Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;

Alignment Scores:
Pred. No.: 1,14e-94 Length: 1876
Score: 1159.50 Matches: 237
Percent Similarity: 60.76% Conservative: 82
Best Local Similarity: 45.14% Mismatches: 115
Query Match: 28.68% Indels: 91
DB: 2 Gaps: 12
```

US-09-303-232-2 (1-770) x AAV12197 (1-1876)

```
QY 297 IleTyrLeuAsnLeuSerAlaLys-----ValCysLeuAlaGlyTyrHisGluLys 313
Db 94 GTCTGGCTGGCGCTGGCCGGTGGCTCTCTGACGTGTCTCCCTGCAAGCGAGTTCAGAGG 153
QY 314 ArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeuAsnGlu 333
Db 154 AAGCTTTACAAGGAGGTGGTCAAGAACTACAATCCCTTGAGAGGCGCGTGGCAATGAC 213
QY 334 SerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAspValAspGlu 353
Db 214 TCGCAACCACTACCGCTCTACTTCTCCCTGAGCTCTCTCGATCATGGACGTGGAGAG 273
QY 354 LysAsnGlnLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeu 373
Db 274 AAGAACAAGTTTAAACCACCAACATTGGCTGCAAAATGCTTGGACAGATCACTATTTA 333
QY 374 ArgTrpAsnThrSerAspTyrGlyValLysAspLeuArgIleProHisArgIle 393
Db 334 CAGTGAATGTGTGAGAATATCCAGGGGTGAAGACTGTTGGTTTCCAGATGGCCAGATT 393
QY 394 TrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGln 413
Db 394 TGGAAACCAAGACATTCTCTATAACAGTGTGATGAGCGCTTTGACGCCACATTCCAC 453
QY 414 ThrAsnValValValArgAsnAsnGlySerCysLeuTyrValProGlyIlePheLys 433
Db 454 ACTAAGCTGTGTGTAATCTTCTGGGCATTGGCCAGTACCTGCTCCAGGCATATTCAAG 513
QY 434 SerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGluMetLys 453
Db 514 AGTTCTGTACATCATGATGATGCTGTTTCCCTTTGATGTGCAGCAGCTGCAAACTGAAG 573
QY 454 PheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAspGluThrGly 473
Db 574 TTTGGGTCTGTGTCTTACGAGCTGTCTTGGATCTGCAGATGCAGAG-----624
QY 474 GlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLys 493
Db 625 GCAGATATCAGTGGCTATATCCCAATGGAGAATGGACCTAGTGGGAATCCCCGCGCAAG 684
QY 494 ArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaIle 513
Db 685 AGGAGTGAAGGTTCATGATGCTGCTCAAGAGAGCCCTACCCCGATGCACCTTCACAGTG 744
QY 514 IleIleArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysValLeuIle 533
Db 745 ACCATCGCGCGGAGGACGCTCTACTATGGCTCAACCTGCTATCCCTGTGTGCTCATC 804
QY 534 AlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeu 553
Db 805 TCGGCCCTCGCCCTGTGTGTGTCTCTGCTCTCTCTGTCAGATCCCGGGGAGAGATTTCCTG 864
```


Db 94 GTCTGGCTGGCGTGGCGGCTGCTCTCTGTGCACGTGTCTCTGTCAAGCGAGTTCACGAGG 153
Qy 314 ArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeuAsnGlu 333
Db 154 AAGCTTTACAGGAGCTGGTCAAGACTACAACTCCCTTGGAGAGCGCGTGGCCAAATGAC 213
Qy 334 SerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnLeuLeuAspValAspGlu 353
Db 214 TCGCAACCACTCACCGTCTACTCTCTCCCTGAGCCTCTCTGCAGATCATGGAGTGGATGAG 273
Qy 354 LysAsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeu 373
Db 274 AAGAACAAGGTTTAAACCAACCAATTTGGCTGCAATGTCTTGGACAGATCACTATTATTA 333
Qy 374 ArgTrpAsnThrSerAspTyrGlyValLysAspLeuArgLeuProProHisArgGlu 393
Db 334 CAGTGGATGTGTGAGATATCCAGGGGTGAAGACTGTCTTCCAGATGGCCAGAT 393
Qy 394 TrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGln 413
Db 394 TGGAAACCAAGACATCTCTCTATAACAGTGTCTGATGAGCGCTTTGACGCCACATCCAC 453
Qy 414 ThrAsnValValAlaArgAsnAsnGlySerCysLeuTyrValProProGlyLeuPheLys 433
Db 454 ACTAACGTGTGTGAATCTCTTGGGCATTGGCAGTACCTGCGCTCCAGGCATATCAAG 513
Qy 434 SerThrCysLysLeuAspLeuThrTrpPheProPheAspAspGlnArgCysGluMetLys 453
Db 514 AGTTCCTGCTACATCGATGACGTGTGTTCCCTTTGATGTCAGCACTGCAAACTGAAG 573
Qy 454 PheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuGlnAspGluThrGly 473
Db 574 TTTGGTCTGTCTTACGGAGGCTGGTCTTGTGATCTGCAGATGCAGGAG----- 624
Qy 474 GlyAspLysSerTrpValLeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLys 493
Db 625 GCAGATATCATGTGCTATATCCCAATGGAGATGGGACCTAGTGGGAATCCCGCGGAAG 684
Qy 494 ArgAsnGluLeuTyrTyrAsnCysCysProGluProTyrLeuAspLeuPheAlaLeu 513
Db 685 AGGAGTGAAGGTTCTATGAGTGTGCAAGAGCCCTACCCCGATGCACCTTCACAGTG 744
Qy 514 IleLeuArgArgThrLeuTyrTyrPhePheAsnLeuLeuLeuProCysValLeuLeu 533
Db 745 ACCATGCGCGCAGACGCTCTACTATGGCCTCAACCTGTGTGATCCCTCTGTGCTCATC 804
Qy 534 AlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeu 553
Db 805 TCCGCCCTCGCCCTGCTGGTGTCTCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 864
Qy 554 GlyValThrLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetPro 573
Db 865 GGGATACAGCTTACTCTCTTACCGTCTTTCATGCTGCTGCTGCTGCTGCTGCTGCTGCTG 924
Qy 574 AlaThrSerAspAlaValProLeu----- 581
Db 925 GCAACATCCGATTCCGTACCATTTGATAGCCAGTACTTTCGCCAGCACCACCATCATCGTG 984
Qy 581 ----- 581
Db 985 GGCCTCTCGTGTGGTGGAGCGGTGATCGTGTGTGAGTACACACACACCGCCGCGAGG 1044
Qy 582 ----- 582
Db 1045 GGCAAGATGCCAAGTGGACAGAGTATCTCTTGAAGTGGTGGCGGTGGTCTCTSCGA 1104
Qy 597 MetSerArgProGly-----ArgProLeuLeuLeuGluPheProThrThrPro 612
Db 1105 ATGAAGAGGCGCCGGGAGGACAGAGTGGCGCCG-----GCCTGCCAGACACAGCGCGCGC 1161
Qy 613 CysSerAspThrSerSerGluArgLysHisGlnLeuLeuSerAspValGluLeuLysGlu 632
Db 1162 TGCAGCTTGGCAGTGTGGAG-----ATGAGCGCGCGTGGCGCGCGCGCGCGCGCGCGCG 1206

Qy 633 ArgSerSerLysSerLeuLeuAlaAsnValLeuAspLeuAspPheArgHis--- 651
Db 1207 GCCAGAACGGGAACCTGTGTGAC-----ATCGGCTTCCGGCGCGCTG 1248
Qy 652 -----AsnCysArgProMetThrProGlyThrLeuProHisAsnProAlaPhe 668
Db 1249 GACGGCGTGCACCTGTGTCCCG---ACCCCC----- 1275
Qy 669 TyrArgThrValTyrGlyGlnGlyAspAspGlySerIleGlyProIleGlySerThrArg 688
Db 1276 -----GACTCTGGGTAGTG-----TGTGGCGCGC 1299
Qy 689 MetProAspAlaValThrHis-----HisThrCysIleLysSerSerThr 703
Db 1300 ATGCCCTGTCTCCCCACGACGATGACACCTCTCTGCACGGCGGGACACCCCGAGGGG 1359
Qy 704 GluTyrGluLeuGlyLeuLeuLeuLysGluLeuArgPheIleThrAspGlnLeuArgLys 723
Db 1360 GACCGGACTTGGCAAAGATCTCTGGAGGAGTCCGCTACATTGCAATCGCTTCCGCTGC 1419
Qy 724 AspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetValValAspArg 743
Db 1420 CAGGACGAAAGCGAGCGGTCTGCAGCGAGTGGAAAGTTCCGCCCTGTGTGGTGGACCGC 1479
Qy 744 LeuCysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeuSer 763
Db 1480 CTGTGCTCTCATGGCTTCTCGGTCTTCACCATCATCTGCACCATCGGCATCTGATGTG 1539
Qy 764 AlaProHisIleIle 768
Db 1540 GCTCCCAACTTCGTG 1554

Search completed: May 8, 2004, 19:29:44
Job time : 850.783 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:31:03 ; Search time 18.738 Seconds
(without alignments)
3952.801 Million cell updates/sec

Title: US-09-303-232-2

Perfect score: 4043

Sequence: 1 MKNAQLKLTVEVDDDELAV.....MFAILLATIAVLISAPHIIVS 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 78:*

1: pir1:*

2: pir2:*

3: pir3:*

4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1182.5	29.2	502	2 A57175	nicotinic acetylch
2	1179.5	29.2	502	2 JN0113	nicotinic acetylch
3	1173.5	29.0	502	2 T01378	nicotinic acetylch
4	1168	28.9	511	2 JH0173	alpha-bungarotoxin
5	1160.5	28.7	502	1 ACHUA7	nicotinic acetylch
6	1155.5	28.6	502	2 G02259	alpha 7 neuronal n
7	1104	27.3	498	2 S68588	nicotinic acetylch
8	1009.5	25.0	461	2 T25671	hypothetical prote
9	988	24.4	560	2 T19622	hypothetical prote
10	968.5	24.0	542	2 T19862	hypothetical prote
11	889	22.0	567	1 ACEFA1	nicotinic acetylch
12	873	21.6	557	2 T12359	nicotinic acetylch
13	845	20.9	503	2 A53956	nicotinic acetylch
14	843.5	20.9	576	1 ACFFA2	nicotinic acetylch
15	841	20.8	521	1 ACFFN2	nicotinic acetylch
16	840	20.8	502	2 A37040	nicotinic acetylch
17	838	20.7	528	1 ACCH2N	nicotinic acetylch
18	828	20.5	511	2 A40110	nicotinic acetylch
19	827.5	20.5	456	1 ACCHAN	nicotinic acetylch
20	824	20.4	499	2 A24572	nicotinic acetylch
21	820.5	20.3	495	2 S60589	acetylcholine rece
22	820.5	20.3	512	2 B37014	nicotinic acetylch
23	817	20.2	494	2 T09289	nicotinic acetylch
24	806.5	19.9	500	2 S12899	nicotinic acetylch
25	805	19.9	498	2 G02421	nicotinic acetylch
26	802.5	19.8	457	1 ACHUA1	nicotinic acetylch
27	801	19.8	457	1 ACBOA1	nicotinic acetylch
28	794.5	19.7	457	2 A28529	nicotinic acetylch
29	794	19.6	457	2 A24383	nicotinic acetylch

ALIGNMENTS

RESULT 1

A57175

nicotinic acetylcholine receptor alpha-7 chain precursor, neuronal - mouse
C:Species: Mus musculus (house mouse)

C:Date: 05-Jan-1996 #sequence_revision 05-Jan-1996 #text_change 20-Aug-1999

C:Accession: A57175

R:Orr-Urtreger, A.; Seldin, M.F.; Baldini, A.; Beaudet, A.L.

Genomics 26, 399-402, 1995

A:Title: Cloning and mapping of the mouse alpha7-neuronal nicotinic acetylcholine recep

A:Reference number: A57175; MUID:95324936; PMID:7601470

A:Accession: A57175

A:Status: preliminary

A:Molecule type: mRNA

A:Residues: 1-502 <ORR>

A:Cross-references: GB:L37663; NID:g790853; PIDN:AAC42053.1; PID:g790854

C:Superfamily: acetylcholine receptor

C:Keywords: brain; glycoprotein; ion channel; neurotransmitter receptor; phosphoprotein

F:1-23/Domain: signal sequence #status predicted <SIG>

F:24-502/Product: nicotinic acetylcholine receptor alpha-7 chain, neuronal #status pred.

F:231-254/Domain: transmembrane #status predicted <TR1>

F:262-280/Domain: transmembrane #status predicted <TR2>

F:296-317/Domain: transmembrane #status predicted <TR3>

F:470-488/Domain: transmembrane #status predicted <TR4>

F:46,90,133/Binding site: carboxylate (Asn) (covalent) #status predicted

F:365,413,427/Binding site: phosphate (Ser) (covalent) #status predicted

F:415/Binding site: phosphate (Thr) (covalent) #status predicted

F:442/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 29.2%; Score 1182.5; DB 2; Length 502;

Best Local Similarity 45.8%; Pred. No. 5e-77;

Matches 239; Conservative 80; Mismatches 118; Indels 85; Gaps 10;

QY 297 IYLNLSA---KVLGAGYHEKRLHLLDLPYNTLERPVNLPVNESDPLQLSGLTLMQIDVDE 353

DB 8 IWLAAALHVSLOGEFQRLYKELVKNYNPLRPVANDSQPLTVYFSLSLQIMDVE 67

QY 354 KNQLLVTVNLKLEWDMNLRNNTSDYGGVKDLRIPPHRIWKFPDVLNMSADEGFGTYQ 413

DB 68 KNOVLTTNWLQMSWTDHYLQNMSEYGVKNVRFPDQGIWKFPDILLYNSADERFDATFH 127

QY 414 TNVVRNNGSLYVPPGIFKSTCKIDITWFFPDQRCMKGSGSWTYDGFQDLQLODRTG 473

DB 128 TNVLNASHGCOYLPFGIFKSCYIDVRWFFPDVQOQKLFKSGWSYGGWSLDLQOE--- 184

QY 474 GDISSYVLNGWELLGVGKNEIYNNCCPEYDITFAIIIRRTLYYFNLIIPCVLI 533

DB 185 ADISSYIPNGWDLMGIFPKRNEKPEYDVTVTMRRRTLYYGLNLLPCVLI 244

QY 534 ASMALLGFTLPDSDGKLSLGVITLLSLTVFLNMVAETMPATSDAVPL----- 581

DB 245 SALALLVFLPADSGEKISLGLTVLLSLTVFLLVAEIMPATSDSVPLIAQYFASTMLIV 304

acetylcholine rece
nicotinic acetylch
nicotinic acetylch
nicotinic acetylch
nicotinic acetylch
acetylcholine rece
nicotinic acetylch
nicotinic acetylch
nicotinic acetylch
protein F25G6.4 [i
nicotinic acetylch
nicotinic acetylch
hypothetical prote
nicotinic acetylch
nicotinic acetylch
hypothetical prote

```
QY 582 -----WIRIVFLCWLPMILRMSRPG-----RPLILEPPTTP 612
Db 305 GLSVVTVIVLXHHHDPDGKMPKXWTRILLNWCWFLRMKPGEDKVRPACQHKPRR- 363
QY 613 CSDTSSERKHOILSDVELKERSKSLANVLDDIDDFRHNCRPMTGGLPHNPAPYRTV 672
Db 364 CS-----LASVELSAGAG-----PPTSGNLLY--IGFRGL 392
QY 673 YQG-----GDDGSI-SPIGSTRMPDVAVTHHTCIKSTVEYELGLILKEIRFITDQLRKDDE 726
Db 393 EGMHCAPTDSGVVVCRLACSPTHDEHLMHGTHSPDGLAKILBEVRYIANRFRQDE 452
QY 727 CNDIANDWKFAAMVVDRLCLLIIFTEFAILLATIAVLSAPHII 768
Db 453 SEVICSEWFAACVVDRLCLMAFSVFTIICIGILMSAPNFV 494

RESULT 2
JN0113
N;Title: A neuronal nicotinic acetylcholine receptor subunit (alpha 7) is developmentally
A;Accession: JN0113
A;Molecule type: DNA
A;Residues: 1-502 <SCH>
A;Cross-references: GB:X68586; NID:G287756; PIDN:CAA48576.1; PID:G287757
A;Experimental source: white leghorn; brain
R;Schoepfer, R.; Conroy, W.G.; Whiting, P.; Gore, M.; Lindstrom, J.
Neuron 5, 35-48, 1990
A;Title: Brain alpha-bungarotoxin binding protein cDNAs and Mabs reveal subtypes of this
A;Reference number: JH0172; MUID:90315158; PMID:2369519
A;Accession: JH0172
A;Molecule type: mRNA
A;Residues: 1-502 <SCH>
A;Cross-references: EMBL:X52295; NID:G63077; PIDN:CAA36543.1; PID:G63078
A;Experimental source: brain
R;Matter-Sadzinski, L.; Hernandez, M.C.; Roztocil, T.; Ballivet, M.; Matter, J.M.
EMBO J. 11, 4529-4538, 1992
A;Title: Neuronal specificity of the alpha7 nicotinic acetylcholine receptor promoter de
A;Reference number: S28018; MUID:93049204; PMID:1425587
A;Accession: S28018
A;Molecule type: DNA
A;Residues: 1-18 <MAS>
A;Cross-references: EMBL:X68246; GB:949751; NID:G65319; PIDN:CAA48317.1; PID:G65320
A;Experimental source: white leghorn; erythrocyte
R;Conti-Tronconi, B.M.; Dunn, S.M.J.; Barnard, E.A.; Dolly, J.O.; Lai, F.A.; Ray, N.; Ra
Proc. Natl. Acad. Sci. U.S.A. 82, 5208-5212, 1985
A;Title: Brain and muscle nicotinic acetylcholine receptors are different but homologous
A;Reference number: A94055; MUID:85270494; PMID:3860855
A;Accession: B25738
A;Molecule type: protein
A;Residues: 24-25, 'Er', 28-41, 'X', 43-45, 'X', 47 <CON>
C;Comment: This acetylcholine receptor is blocked by alpha-bungarotoxin and is localized
C;Genetics:
A;Introns: 19/1; 65/3; 80/3; 117/2; 144/1; 200/1; 265/1; 294/1; 330/3
C;Superfamily: acetylcholine receptor
C;Keywords: brain; glycoprotein; ion channel; neurotransmitter receptor; phosphoprotein;
F;1-23/Domain: signal sequence #status predicted <SIG>
F;24-502/Product: nicotinic acetylcholine receptor alpha-7 chain, neuronal #status predi
F;231-254/Domain: transmembrane #status predicted <TR1>
F;262-280/Domain: transmembrane #status predicted <TR2>
F;296-317/Domain: transmembrane #status predicted <TR3>
F;470-488/Domain: transmembrane #status predicted <TR4>
F;46, 90, 133/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;365, 367, 413, 427, 465/Binding site: phosphate (Ser) (covalent) #status predicted
F;445/Binding site: phosphate (Thr) (covalent) #status predicted
F;442/Binding site: phosphate (Tyr) (covalent) #status predicted
```

```
Query Match 29.2%; Score 1179.5; DB 2; Length 502;
Best Local Similarity 45.6%; Pred. No. 8.2e-77;
Matches 237; Conservative 80; Mismatches 127; Indels 75; Gaps 9;

QY 295 LIYINLSA---KVCLAGYHEKRLHDLDPYNTLIRPVLNBSDDPQLQSLGTLMOIDV 351
Db 6 LMLLILAAAGLVRESLQGEFQKLYKELLKNTNPLRPVANDSQPLTVYFTLSLMOIMDV 65
QY 352 DEKNQLVTNVWLKLEWDMNLRWNTSDYGGVKDLRIIPHRITWKPDVLMYNSADEGPDGT 411
Db 66 DEKNQVLTNIWLQYWTWTHYLQWNVSEYPGVKNVRFPPDGLIWKPDILLYNSADERDAT 125
QY 412 YOTNVVRNNGSLYVPPGIFKSTCKIDITWFFDDQRCMEKFGSTYDQFQLDLQLQDE 471
Db 126 FHTNVLVNSGHCQYLPFGIFKSSCYIDVRWFFDVQKCNLKFSGTYGWSLDLQME- 184
QY 472 TGGDISYVINGEWELLVPGVKRNEIYNCCPPYDITFAIIRRTLYYFNLIIPCV 531
Db 185 --ADISGYISNGEWDLVGIPGKRTESFYECCKEPYDITFTVTRRTLYYGLNLIPCV 242
QY 532 LIASMLLGLTLPDPSGEKLSLGTILSLTVFLNVAETMPATSDAVPL----- 581
Db 243 LISALALLVFLPADSGEKLSLGTIVLLSLTVFMLLVAEIMPAISVPLAQYFASTMI 302
QY 582 -----WIRIVFLCWLPMILRMSRPGRPLILEPPTTPCS 614
Db 303 IVGLSVVTVIVLQYHHHDPDGKMPKXWTRIVLLNWCWFLRMKRG-EDKVRPAC 357
QY 615 DTSSERKHOILSDVELKERSKSL-ANVLDDIDDFRH-----NCRPMTGGLPHNPAPY 669
Db 358 QHKQRCS--LSSMEMNTVSGQCSNGNMLY--GFRGLDGVRCTPTDTSGLV----- 406
QY 670 RTVYGGDDSGIGPIGSTRMPDVAVTHHTCIKSTVEYELGLILKEIRFITDQLRKDDE 729
Db 407 -----CGRMTCSPTENLHSGPSPDLAKILBEVRYIANRFRQDEEEA 455
QY 730 IANDWKFAAMVVDRLCLLIIFTEFAILLATIAVLSAPHII 768
Db 456 ICNEWKAASVVDRLCLMAFSVFTIICIGILMSAPNFV 494

RESULT 3
T01378
nicotinic receptor alpha 7 chain - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 12-Feb-1999 #sequence_revision 12-Feb-1999 #text_change 21-Jul-2000
C;Accession: T01378
R;Seguela, P.; Wadiche, J.; Dineley-Miller, K.; Dani, J.A.; Patrick, J.W.
J. Neurosci. 13, 596-604, 1993
A;Title: Molecular cloning, functional properties, and distribution of rat brain alpha
A;Reference number: Z14310; MUID:93147931; PMID:7678857
A;Accession: T01378
A;Status: preliminary; translated from GB/EMBL/DBJ
A;Molecule type: mRNA
A;Residues: 1-502 <SEG>
A;Cross-references: EMBL:S53987; NID:G264770; PIDN:AAB25224.2; PID:G5705903
A;Experimental source: brain
C;Superfamily: acetylcholine receptor

Query Match 29.0%; Score 1173.5; DB 2; Length 502;
Best Local Similarity 45.6%; Pred. No. 2.2e-76;
Matches 238; Conservative 80; Mismatches 119; Indels 85; Gaps 10;

QY 297 IYINLSA---KVCLAGYHEKRLHDLDPYNTLIRPVLNBSDDPQLQSLGTLMOIDVDE 353
Db 8 IWLALAAALHVLQGEFQKLYKELLKNTNPLRPVANDSQPLTVYFTLSLQIMDVDE 67
QY 354 KNQLVTNVWLKLEWDMNLRWNTSDYGGVKDLRIIPHRITWKPDVLMYNSADEGFDGTQ 413
Db 68 KNQVLTNIWLQWSTWTHYLQWNVSEYPGVKNVRFPPDGLIWKPDILLYNSADERDATFH 127
QY 414 TNVVRNNGSLYVPPGIFKSTCKIDITWFFDDQRCMEKFGSTYDQFQLDLQLQDETG 473
```



```

Db 128 TNLVNASGHQCLPGLPKSSCYIDVRWFPDVOQCKLFGSWSYGGSLDLQMBE---184
QY 474 GDSSVVLNGEWELLGVPGKRNELYYECKEPEYDVTYITMRRRTYYGL 243
Db 185 ADISSYIPNGEWMIPGKRNELYYECKEPEYDVTYITMRRRTYYGLNLLIPCVLI 244
QY 534 ASMALGFTLPDPSGKLSIGVTLLSLVFLNVAETMPATSDAVPL---581
Db 245 SALALLVFLPADSGEKISGLITVLLSLVFLNVAETMPATSDVPLIAQVASTMIIV 304
QY 582 -----WIRIVFLCWLPLWILMSRPG-----RPLILEPPTP 612
Db 305 GLSVVVTVIVLRVHHHDPDGKMPKXTRIILLNKCWFLRMKEPGEDKVRPACQKPRR- 363
QY 613 CSDTSERKHQILSDVELKERSKSLIANVLDDDDFRHNCRPMTPGGTLPHNPAFYRTV 672
Db 364 CS-----LASVELSAGAG-----PPTSGNLLY--IGFRGL 392
QY 673 YGQ-----GDGSI-GPIGSTRMPDAVTHHTCTKSSTEYELGLILKEIRFITDOLRQDE 726
Db 393 EGHICAPTDSGVVCGRLACSPTHDHLMHGAPSDGDPDLAKILEBRYIANRRCODE 452
QY 727 CNDIANDKFAAMVVDRLCLIIFTMFALATIIVALLSAPHII 768
Db 453 SEVICSEWKFACVVDPLCLMAFSVFTICTIGILMSAPNFV 494

RESULT 4
JH0173
A:Title: alpha-bungarotoxin-binding protein alpha-2 chain precursor - chicken
A:Species: Gallus gallus (chicken)
A:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 20-Aug-1999
A:Accession: JH0173
R:Schoeffer, R.; Conroy, W.G.; Whiting, P.; Gore, M.; Lindstrom, J.
Neuron 5, 35-48, 1990
A:Title: Brain alpha-bungarotoxin binding protein cDNAs and MABs reveal subtypes of this
A:Reference number: JH0172; MUID:90315158; PMID:2369519
A:Accession: JH0173
A:Molecule type: mRNA
A:Residues: 1-511 <SCH>
A:Cross-references: GB:X52296; NID:G63081; PIDN:CAA36544.1; PID:G63082
A:Experimental source: brain
A:Note: this sequence is similar to acetylcholine receptor alpha chains
C:Comment: Alpha-bungarotoxin binding proteins are localized to extrasynaptic pseudodend
C:Superfamily: acetylcholine receptor
C:Keywords: glycoprotein; transmembrane protein
F:1-30/Domain: signal sequence #status predicted <SIG>
F:31-510/Product: alpha-bungarotoxin binding protein alpha-2 chain #status predicted <AR
F:239-262/Domain: transmembrane #status predicted <TM1>
F:270-288/Domain: transmembrane #status predicted <TM2>
F:304-323/Domain: transmembrane #status predicted <TM3>
F:479-496/Domain: transmembrane #status predicted <TM4>
F:547/Binding site: carbohydrate (Asn) (covalent) #status predicted

Query Match 28.9%; Score 1168; DB 2; Length 511;
Best Local Similarity 44.2%; Pred. No. 5.6e-76;
Matches 238; Conservative 83; Mismatches 117; Indels 100; Gaps 12;

QY 285 NGLKHSWIFLLIYLNLSAKVCLAGYVHEKRLHLDLPYNTLPERPVNLSDEPLQLSFGIT 344
Db 12 SGLCLWASLFLGFF-----KVSCQGESQRRLYRDLRLNRYNLERPWNQSDQIVVELQLS 66
QY 345 LMQILDVDEKQQLLVTVNWVLEWDMNLRWNTSDYGVKVDRIPIPHRIWKPDVLMYNSA 404
Db 67 LQQLIDVDEKQQLITNAWLOWYWDIYLSVDQYEVGVQNLRFPSDQIWIWPDILLXNSA 126
QY 405 DSGFGDTGYTVNVNRNNGSLVVPBGIFKSTCKIDITWPPDDQRCMKFGSWTYDGLQ 464
Db 127 DERFATHTNVLVNYSVGSQVIPPGLIKSTCYIDVRWFPDVOQCKDLFGSWTHSGMWI 186
QY 465 DILOQDETGGDISSVLNGEWELLGVPGKRNELYYECKEPEYDVTYITMRRRTYYGL 524

```

```

Db 187 DLQMLE---ADISNYLSNGEWLLGVPGKRNELYYECKEPEYDVTYITMRRRTYYGL 243
QY 525 NLIIPCVLIASMALGFTLPDPSGKLSIGVTLLSLVFLNVAETMPATSDAVPL---581
Db 244 NLIIPCVLIASMALGFTLPDPSGKLSIGVTLLSLVFLNVAETMPATSDVPLIAQ 303
QY 582 -----WIRIVFLCWLPLWILMSRPG-----RPL 604
Db 304 YFASIMVIVGLSVVVTVLVLQFHHDPPQAGKMPRVVRVILLNKCWFLRMKKGKGENIKPL 363
QY 605 ILEFPTTSCDTSERKHQILSDVELKERSKSLIANVLDDDDFRHNCRPMTPGGTLPH 664
Db 364 SCKY-----SYPKHPSLKNTEM-----NVL-----PG-----H 387
QY 665 NPAFVTVTVG-----QGDD--GSIGPIGSTRMPDAVTHHTCTKS--STEYELGLI 710
Db 388 QPSNGNMISYHTMTMENPCPPQNDLGSKGKITCPDSED--NEHVQKKALMDTIPVVKI 445
QY 711 LKEIRFITDOLRQDEKDDCEDNDIANDKFAAMVVDRLCLIIFTMFALATIIVALLSAPHII 768
Db 446 LEEVOFIAMRFRKQDEGEIEICSEWKFAPAAVIDRLCLVAFTLFAICTFTILMSAPNFI 503

RESULT 5
ACHUA7
nicotinic acetylcholine receptor alpha-7 chain precursor, neuronal - human
N:Alternate names: cholinergic nicotinate receptor alpha-7 chain
C:Species: Homo sapiens (man)
C:Date: 12-Aug-1996 #sequence_revision 31-Jan-1997 #text_change 22-Jun-1999
A:Accession: I37185; A54194; S60309
R:Peng, X.; Katz, M.; Gerzanich, V.; Anand, R.; Lindstrom, J.
Mol. Pharmacol. 45, 546-554, 1994
A:Title: Human alpha 7 acetylcholine receptor: cloning of the alpha 7 subunit from the
omers expressed in Xenopus oocytes.
A:Reference number: I37185; MUID:94195283; PMID:8145738
A:Accession: I37185
A>Status: translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-502 <PEN>
A:Cross-references: EMBL:X70297; NID:G496606; PIDN:CAA49778.1; PID:G496607
A:Experimental source: brain neuroblastoma cell line SHSY-5Y
R:Chini, B.; Raimond, E.; Elgoyhen, A.B.; Moralli, D.; Balzaret, M.; Heinemann, S.
Genomics 19, 379-381, 1994
A:Title: Molecular cloning and chromosomal localization of the human alpha-7-nicotinic :
A:Reference number: A54194; MUID:94245214; PMID:8188270
A:Accession: A54194
A:Molecule type: mRNA
A:Residues: 24-363, 'S', 365-374, 'A', 376-408, 'AWPAP', 414-502 <CHI>
A:Cross-references: GB:Z23143; NID:G457736; PIDN:CAA80672.1; PID:G457737
A:Experimental source: retina
C:Comment: This acetylcholine receptor is blocked by alpha-bungarotoxin and is localized
C:Genetics:
A:Gene: GDB:CHRNA7
A:Cross-references: GDB:138751; OMIM:118511
A:Map position: 15q14-15q14
A:Note: defects in this gene have been associated with mental retardation and schizophre
C:Complex: the functional receptor molecule is a heteropentamer with two alpha chains ar
C:Superfamily: acetylcholine receptor
C:Keywords: brain; glycoprotein; heteropentamer; ion channel; neurotransmitter receptor;
F:1-23/Domain: signal sequence #status predicted <SIG>
F:24-502/Product: nicotinic acetylcholine receptor alpha-7 chain, neuronal #status predi
F:231-254/Domain: transmembrane #status predicted <TR1>
F:262-280/Domain: transmembrane #status predicted <TR2>
F:296-317/Domain: transmembrane #status predicted <TR3>
F:470-488/Domain: transmembrane #status predicted <TR4>
F:46.90.133/Binding site: carbohydrate (Asn) (covalent) #status predicted
F:150-164/Disulfide bonds: #status predicted
F:365.413/Binding site: phosphate (Ser) (covalent) #status predicted
F:415/Binding site: phosphate (Thr) (covalent) #status predicted
F:442/Binding site: phosphate (Tyr) (covalent) #status predicted

Query Match 28.7%; Score 1160.5; DB 1; Length 502;
Best Local Similarity 45.1%; Pred. No. 1.9e-75;

```

[illegible]

RESULT 6
G02259
alpha 7 neuronal nicotinic acetylcholine receptor - human
C:Species: Homo sapiens (man)
C:Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 20-Aug-1999
C:Accession: G02259
R:Leonard, S.
submitted to the EMBL Data Library, November 1995
A:Reference number: H00936
A:Accession: G02259
A:Status: preliminary; translated from GB/EMBL/DBDJ
A:Molecule type: mRNA
A:Residues: 1-502 <L80>
A:Cross-references: EMBL:U40593; NID:g1125076; PID:g1125077
C:Superfamily: acetylcholine receptor

Query Match	28.6%;	Score 1155.5;	DB 2;	Length 502;
Best Local Similarity	45.1%;	Pred. No. 4.3e-75;		
Matches 237;	Conservative 80;	Mismatches 117;	Indels 91;	Gaps 12;
QY	297	IYLNLSAK--VCLAGVHEKRILLDDLPYNTLERPVLNESDPQLQSLFGLTLMQIIDVDE	353	
Db	8	VWLAALASLLHVSLLQGEFQRKLYKELVKYNPFLERPANDSQPLTVYFSLNLLQIMDVDE	67	
QY	354	KNQLLVTNVWLKLEWNIMLWNTSYDGGVKDLRPPHRIWKPVLWYNSADEGFDGYQ	413	
Db	68	KNQVLTINWLQMSWTHYLNWVSEYGVKTVTFPDQGIWKPDIILLVNSADERDFATFH	127	
QY	414	TNVVNRNGSCLYVPPGFKSTCKIDIITWFFPDQORCEMKPGSWTYDGFQDLQLQDSTG	473	
Db	128	TNVLVNPSGHCOYLPPIFKSCYIDVAFWFFDQHQKLFKGSWSYGGWSDLQMQE---	184	
QY	474	GDISSVILNGEWELGVFGKRNIEYINCCPPFYDITFAIIRRTLYFFYNLIIPCVLII	533	

```

Db      185 ADISGYTPNGEWDLVGIPGRKSERFECCKEPDPDVTFVTVMRRRTLIYGINLLIPCVI 244
Qy      534 ASMALLGFTLPDPSGEKLSGVTLISLTVFLNNVAETMPATSDAVPL-----581
       ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db      245 SALALLVFLPADSGEKLSLIGITVLLSLTVFMLLVAEIMPATSDSVPLIAQYFASTMIIV 304
Qy      582 -----WTRIVFLCWLPIWLEMSRPG---RPLILEPPTP 612
Db      305 GLSVVVTVIVLQXHHDDPDGGKKPKTRVILLNWCWAFLEMKRPGEDKVVP-ACQHKQR 363
Qy      613 CSDTSSERKHQILSDVELKERSSKLIANVLIDDDFRH---NCRPMTEGGTIPHNPAPF 668
Db      364 CSLASVE----MSAVAPPASPANGNLLY--IGFRGLDGVCVP-TP-----401
Qy      669 YRTVYGOGDGSIGTGSTMPDAVTH-----HTCTKSSTEYEELGLIKBEIRITDLRK 723
Db      402 -----DSGVV---CGSMAGSPTHDEHLLHGGOPEGDPLAKILEEVRIANRFRC 449
Qy      724 DDECNDIANDWKFAAMVVDRLCLIIITMFAILATIIVLLSAPHII 768
Db      450 QDESEAVECWKPAAVCVVDRLCLMAFSVFTHCTIGILMSAPNFV 494

RESULT 7
S68588
nicotinic acetylcholine receptor alpha-1 chain precursor (clone Ce21) - Caenorhabditis
C;Species: Caenorhabditis elegans.
C;Date: 06-Dec-1996 #sequence_revision 07-Feb-1997 #text_change 20-Aug-1999
C;Accession: S68588; S57496
R;Ballivet, M.; Alliod, C.; Bertrand, S.; Bertrand, D.
J. Mol. Biol. 258, 261-269, 1996
A;Title: Nicotinic acetylcholine receptors in the nematode Caenorhabditis elegans.
A;Reference number: S68587; MUID:96196478; PMID:8627624
A;Accession: S68588
A;Status: nucleic acid sequence not shown
A;Molecule type: mRNA
A;Residues: 1-498 <BAL>
A;Cross-references: EMBL:X83887; NID:g872087; PIDN:CAA58764.1; PID:g872088
C;Superfamily: acetylcholine receptor
C;Keywords: glycoprotein; ion channel, neurotransmitter receptor; postsynaptic membrane.
F;1-19/Domain: signal sequence #status predicted <SIG>
F;20-498/Product: nicotinic acetylcholine receptor alpha-1 chain #status predicted <MAT>

Query Match          27.3%; Score 1104; DB 2; Length 498;
Best Local Similarity 43.3%; Pred.No.2.1e-71;
Matches 227; Conservative 76; Mismatches 139; Indels 82; Gaps 8;

Qy      295 LLIVNLNSAKVCLAGYHEKEKLHLHDLDPPNTLRPVINESDPLQLSFGTLMQIIDVDEK 354
Db      6 LLISALIAAPTJGSLOERLYDELKNYNLNLRVANHSPEVTVHLKVALQQIIDVDEK 65
Qy      355 NQLLIVTNVWLKLEWMDNMNRWNTSDYGGVKDLRIHPHRWKPDVLMYNSADEGFGTYQT 414
Db      66 NVVVVNADLDYTWNVDYLVWDKAEGYNIQTDVRFPAGKIWKPDVLLYNSVDTFDSTYQT 125
Qy      415 NVVVNNGSCLYVPPGIFKSTCKIDITWFDFDORCEMKFGSWTYDGFOLDQLQDSGTGG 474
Db      126 NMIVYSTGLVHWVPPGIFKISKIDIQWFFDEQKCFKFGSWTYDGYKLD--LQPATGG 183
Qy      475 DISSVYLNGEWELLGVPGKRNEYNYCCPEPYIDIITFAIIIRRRTIYYFNLIIPCULI 533
Db      184 FDISEYINGEWALPLITVERNEKFDCPEPPDVHFVFLHMRRRTLIYGNLIMPCILT 243
Qy      534 ASMALLGFTLPDPSGEKLSLGVTLISLTVFLNNVAETMPATSDAVPL-----581
Db      244 TMTLLGFTLPDPAAGEKITLQITVLLSICFELSIVSEMSPTSEAVPLLGIFFTCMVIV 303
Qy      582 -----WTRIVFLCWLPIWLEMSRPGRLILFEFPTPCSDT 616
Db      304 TASTVTVTVVLNHYRTPETHDMGPWTRNLLTWIPWILRMKRGHN--ITYASLPSLFS 361
Qy      617 SSERKHQILDVELKERSSSKLIANVLID-----IDDFRHNCRMPPTGGTLPHNP 666
       ::::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```


R;Hembry, C.
 submitted to the EMBL Data Library, March 1996
 A:Reference number: Z19188
 A:Accession: T19862
 A>Status: preliminary; translated from GB/EMBL/DDBJ
 A:Molecule type: DNA
 A:Residues: 1-542 <WIL>
 A:Cross-references: EMBL:Z70266; PIDN:CAA94206.1; GSPDB:GN00028; CBSP:C40C9.2
 A:Experimental source: clone C40C9
 C:Genetics:
 A:Gene: CBSP:C40C9.2
 A:Map position: X
 A:Introns: 11/3; 69/3; 123/2; 173/3; 216/1; 248/1; 279/1; 314/1; 350/3; 430/1; 464/1; 49
 C:Superfamily: acetylcholine receptor

Query Match 24.0%; Score 968.5; DB 2; Length 542;
 Best Local Similarity 38.7%; Pred. No. 1.2e-61;
 Matches 208; Conservative 79; Mismatches 144; Indels 107; Gaps 12;

QY 312 EKRLHLLDPVNTLRRPVLNESDPLQLSFGTLTMOIIDVDEKNQLLVNVLKLEWDM 371
 Db 30 EYRLADLRHNDPYPERPANASEPLVSVKIVLQILDVDEKNQVTLVAMLEYQWTDY 89
 QY 372 NIWRNTSDYGGKDRIP--PHRIWKPDVLMYNSADEGFGDTYQTNVVRNNGSCLYVPP 429
 Db 90 KLKWDPSYGGIKDIRPGNAINAKPDVLLYNSADENFDSTYPVNVVSYTGDLQVPP 149
 QY 430 GIFKSTCKIDITWFPDDQRCCKMFGSWTYDGFQDLQLO-----DETGGDISYVLN 482
 Db 150 GILKUSCKIDITWFPDDQICHLKFGSWTYSGNFIDLRINGPGRKNISDEGIDVQYVQN 209
 QY 483 GEWELGVGKNEIYNNCCPEYIDITFAIIIRRRITLYFENLIIPCVLIASMLLGT 542
 Db 210 GEWLLAVARHETNFD--EQYPSLFFVLIIRLTLYGLNLIIPSLISLMTVLGT 267
 QY 543 LPDSEKSLGVTLLSITVFLNMVAETMPATSDAVPL----- 581
 Db 268 LPDAGEKITLEITILLSCVFFLSWADMTPTSEAVPLTGLIFSGAFFSCMLVVSAS 327
 QY 582 -----WIRIVLCWLPWILMRSPGRPLILEPTTPCSDTSSR 619
 Db 328 VVFTVLVNLHNRKPEHTEMSPLRELLIWLPLLMRRPGKITP-----NCHLKAE 381
 QY 620 RKHQLSDVELKE-----RSSKSLIANV-----LDIDDDF--RHN-CRPM 656
 Db 382 KAEKAKQGSIRKNGVGPKEPDSVHPSEGLSLMKNIKLGRQQIDFEYEFHVQHNLMPV 441
 QY 657 TPGGTLPHNPAFRTVYVGQDDGSGIGPISGTRMPDAVTHHTCTKSSTY-----ELGLI 710
 Db 442 APSEMTF-----RVTY-----SKVMAESYVEDVVMTELKYNMQKACLEKNI 483
 QY 711 LKEIRFITQLRKDDCENDIANDKFAAMVVDRLCLIFTMEAILATIALLSAPHI 768
 Db 484 SSGTRAMRKMEEDERDEQAANDKFAAMVVDRLCLITFSVFLVSTCGIMFSPPLHI 541

RESULT 11
 ACFFAI
 nicotinic acetylcholine receptor alpha-1 chain precursor - fruit fly (Drosophila melanogaster)
 C:Species: Drosophila melanogaster
 C:Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 21-Jun-2002
 C:Accession: S00381; A38801
 R;Bossy, B.; Ballivet, M.; Spierer, P.
 EMO J. 7, 611-618, 1988
 A:Title: Conservation of neural nicotinic acetylcholine receptors from Drosophila to ver
 A:Reference number: S00381; MUID:88283626; PMID:2840281
 A:Accession: S00381
 A:Molecule type: DNA
 A:Residues: 1-567 <BOS>
 A:Cross-references: GB:X07194; NID:g7575; PIDN:CAA30172.1; PID:g7576
 A:Accession: A38801
 A:Molecule type: mRNA
 A:Residues: 1-567 <BOS>

A:Cross-references: EMBL:X07194; NID:g7575; PIDN:CAA30172.1; PID:g7576
 A:Note: 538-Tyr was also found
 C:Genetics:
 A:Gene: FlyBase:nAcR-alpha-96Aa
 A:Cross-references: FlyBase:FBgn0000036
 A:Map position: 3R 96A
 A:Introns: 64/3; 79/3; 116/2; 176/3; 330/2; 401/1; 499/3
 C:Superfamily: acetylcholine receptor
 C:Keywords: glycoprotein; ion channel; neurotransmitter receptor; postsynaptic membrane;
 F;1-21/Domain: signal sequence #status predicted <SIG>
 F;22-567/Product: nicotinic acetylcholine receptor alpha-like chain #status predicted <W>
 F;22-240/Domain: extracellular #status predicted <EXT>
 F;240-264/Domain: transmembrane #status predicted <TM1>
 F;272-290/Domain: transmembrane #status predicted <TM2>
 F;306-325/Domain: transmembrane #status predicted <TM3>
 F;326-513/Domain: intracellular #status predicted <INT>
 F;514-532/Domain: transmembrane #status predicted <TM4>
 F;45,233/Binding site: carboxylate (Asn) (covalent) #status predicted
 F;149-163,222-223/Disulfide bonds: #status predicted

Query Match 22.0%; Score 889; DB 1; Length 567;
 Best Local Similarity 35.2%; Pred. No. 6.6e-56;
 Matches 196; Conservative 91; Mismatches 168; Indels 102; Gaps 10;

QY 291 SWIFLLIYLNLS-AKVCLAGYHEKRLHLLDLPYNTLRRPVLNESDPLQLSFGTLTMOII 349
 Db 3 SVLFAAVFIALHFATGGLANPDARKLYDLDLSNYNLIRPVGNSDRLTVKMSLRLSQLI 62
 QY 350 DVDEKNQLLVNVLKLEWDMNLRWNTSDYGGKDLRIPPHRIWKPDVLMYNSADEG 409
 Db 63 DVNLKNQIMTINWVEQEWNDYKLKNPDDYGGVDTLHPVSEHIMHPDVLVYNNADGNVE 122
 QY 410 GYQTNVVRNNGSCLYVPPGIFKSTCKIDITWFPDDQRCCKMFGSWTYDGFQDL 466
 Db 123 VTIMTKAIIHHGKVKVWKPEPAIYKSFCEIDVEYFPDEQTCFMKFGSWTYDGMVDRL 182
 QY 467 -QLQD-----ETGDISSYVLNGEWELLGVPGRNEIYNNCCPEYIDITFAIIIRRTLY 521
 Db 183 KQTADSDNIEVGIDLDQYIYSVEWDIMRVPAVRNEKFXSCPEYIDIVFNLTLRKTLF 242
 QY 522 YFENLIIPCVLTASMLLGFITPPDSGEKLSLGTITLLSITVFLNMVAETMPATSDAVPL 581
 Db 243 YTNLIIPCVGISFSLVFLYLPDSGEKLSICISILLSTLVFFLLALAIIPSTVTVPL 302
 QY 582 -----WIRIVLCWLPWILMRSPGRPL 604
 Db 303 LGKYLFTMVLVTLVVVTVIIVNFRSPVTHRMAPVQRLFIQLPKLICIERPKK-- 360
 QY 605 ILEFPPTPCSDTSSERKHQILSDV-ELKERSKSLIANVLIDDDDFRHCNCRMPGGLP 663
 Db 361 -----EPEEDQPEVLTDVHLPP-----DVKFVNYDSKRFSGDYGIP 400
 QY 664 HNPAYVR-----TVYQGDD-----GSTGPGTSTEMPDAVTH 695
 Db 401 ALPASHRFDLAAAGSIAHCFAPPLPSSLPDPGADDDLFSPSGLNGDISPGCCPAAAA 460
 QY 696 HTCIKSST-----EYELGLILKEIRFITQLRKDDCENDIANDKFAAMVVDRLCLIFT 750
 Db 461 AAADLSPTPEKPYAREMEXTIEGSRFIAQHVKNKDKFESVEDKQYVAMVLDRLMELWIFA 520
 QY 751 MFAILATIALLSAPHI 767
 Db 521 IACVVGITALIILQAPSL 537

RESULT 12
 S12359
 nicotinic acetylcholine receptor alpha-1 chain precursor - desert locust
 C:Species: Schistocerca gregaria (desert locust)
 C:Date: 31-Dec-1993 #sequence_revision 31-Dec-1993 #text_change 20-Aug-1999
 C:Accession: S12359
 R;Marshall, J.; Buckingham, S.D.; Shingai, R.; Lunt, G.G.; Goosey, M.W.; Darlison, M.G.;
 EMO J. 9, 4391-4396, 1990


```

Db      187 NKNFVLDSDYKSGTWDIEVPAYLN-VYEGDSNHPETDITFYIIIRKRTFYVNLIL 245
QY      529 PCVLIASMAILGFTLPDPSGKLSLGVITLLSLTVFLNMVAETMPATS----- 576
Db      246 PTVLISFLCVLFYLPAAEAGEKVTLGISILLSLVFLLVSKILPPTSLVPLIAKYLLF 305
QY      577 -----DAVPLWIRIVFLCWLPLWILKSRPGRPLILEFPTT 611
Db      306 TFMNTVSIIVTVIINWNERGPRTHRMFMYIRISIFLHYLPAPLFMKRPRKTRLRWMMEM 365
QY      612 PCSDTSSERKHQILSDVELKERS-----KSLLANVLIDDDFRHNC---RPMTPGGTLPH 664
Db      366 PGMSMPAHPHPSYSGPAELPKHISAIGGKQSKWEMVELSDLHHPNCKINEKVNSSGEL-- 423
QY      665 NPAFYTVVYQGDGSGIGPIGSTRMPDPAVTHHTCIKSSTEYELGLLILKEIRFITDOLRKD 724
Db      424 -----GLGD---GCRSESSDSILLSPEASKATE-----AVEFTAEHLRNE 462
QY      725 DECNDIANDWKFAAMVVDRCLLIIFTMFAILATIAVLISAPHI 767
Db      463 DLYIQTRDQKYAMVIDRLQLYIFFIVTTAGIVGILMDAPHI 505

```

Search completed: May 7, 2004, 11:39:30
 Job time : 20.738 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:18:28 ; Search time 14.3803 Seconds

(without alignments)
2788.120 Million cell updates/sec

Title: US-09-303-232-2

Perfect score: 4043

Sequence: 1 MKNAQLKLTVDDELWLAV.....MPAILATIAVLSSAPHIIVS 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt_42.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1182.5	29.2	502	1 ACH7 MOUSE	P49582 mus musculus
2	1179.5	29.2	502	1 ACH7 CHICK	P22770 gallus galli
3	1173.5	29.0	502	1 ACH7 RAT	Q05941 rattus norv
4	1159.5	28.7	502	1 ACH7 HUMAN	P36544 homo sapien
5	1157	28.6	498	1 ACH7 BOVIN	P54131 bos taurus
6	1104	27.3	438	1 ACH7 CAEEL	P48180 caenorhabdi
7	888	22.0	567	1 ACH1 DROME	P09478 drosophila
8	886	21.9	516	1 ACH1 MANSE	P91766 manduca sex
9	873	21.6	557	1 ACH1 SCHGR	P23414 schistocerc
10	848.5	21.0	529	1 ACH2 HUMAN	Q15822 homo sapien
11	845	20.9	503	1 ACH3 HUMAN	P32237 homo sapien
12	843.5	20.9	576	1 ACH2 DROME	P17644 drosophila
13	840	20.8	521	1 ACH3 DROME	P04755 drosophila
14	838	20.7	528	1 ACH2 CHICK	P09480 gallus galli
15	827.5	20.5	456	1 ACHA CHICK	P09479 gallus galli
16	827.5	20.5	496	1 ACH3 CHICK	P09481 gallus galli
17	827	20.5	511	1 ACH2 RAT	P12389 rattus norv
18	823	20.4	499	1 ACH3 RAT	P04757 rattus norv
19	820.5	20.3	495	1 ACH3 BOVIN	Q07263 bos taurus
20	820.5	20.3	512	1 ACH3 CARAU	P18845 carassius a
21	815.5	20.2	519	1 ACH4 DROME	P25162 drosophila
22	805	19.9	498	1 ACHP HUMAN	P30926 homo sapien
23	803	19.9	494	1 ACH6 CHICK	P49581 gallus galli
24	801	19.8	457	1 ACHA BOVIN	P02709 bos taurus
25	799	19.8	457	1 ACH6 HUMAN	Q15825 homo sapien
26	797	19.7	454	1 ACHA MOUSE	P04756 mus musculus
27	794.5	19.7	457	1 ACH2 XENLA	P05377 xenopus lae
28	792.5	19.6	493	1 ACH6 RAT	P43143 rattus norv
29	791.5	19.6	538	1 ACH8 CAEEL	P45963 caenorhabdi
30	789	19.5	457	1 ACHA RAT	P25108 rattus norv
31	789	19.5	470	1 ACHP CHICK	P26153 gallus galli
32	787	19.5	491	1 ACHN CHICK	P09484 gallus galli
33	787	19.5	629	1 ACH4 MOUSE	O70174 mus musculus

34	786.5	19.5	456	1 ACHA_BRARE	Q98880 brachydanio
35	785	19.4	461	1 ACHA_TORMA	P02711 torpado mar
36	783	19.4	495	1 ACHP RAT	P12392 rattus norv
37	782.5	19.4	500	1 ACHN RAT	P12390 rattus norv
38	782	19.3	502	1 ACHN_HUMAN	P17787 homo sapien
39	780	19.3	482	1 ACHA_HUMAN	P02708 homo sapien
40	780	19.3	627	1 ACHA_HUMAN	P43681 homo sapien
41	779.5	19.3	622	1 ACH4_CHICK	P09482 gallus galli
42	779	19.3	461	1 ACHA_TORCA	P02710 torpado cal
43	773.5	19.1	457	1 ACH1_XENLA	P22456 xenopus lae
44	772	19.1	464	1 ACHO RAT	P12391 rattus norv
45	769.5	19.0	459	1 ACHN_CARAU	P19370 carassius a

ALIGNMENTS

RESULT 1

ID	ACH7 MOUSE	STANDARD;	PRT;	502 AA.
AC	P49582;			
DT	01-FEB-1996 (Rel. 33, Created)			
DT	01-FEB-1996 (Rel. 33, Last sequence update)			
DT	28-FEB-2003 (Rel. 41, Last annotation update)			
DE	Neuronal acetylcholine receptor protein, alpha-7 chain precursor.			
GN	CHRNA7 OR ACR47.			
OS	Mus musculus (Mouse).			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.			
OX	NCBI_TaxID=10090;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RC	STRAIN=BALB/c; TISSUE=Brain;			
RX	MEDLINE=95324936; PubMed=7601470;			
RA	Orr-Urtreger A., Seldin M.F., Baldini A., Beaudet A.L.;			
RT	"Cloning and mapping of the mouse alpha 7-neuronal nicotinic acetylcholine receptor."			
RT	Genomics 26:399-402(1995).			
CC	-!- FUNCTION: After binding acetylcholine, the AChR responds by an extensive change in conformation that affects all subunits and leads to opening of an ion-conducting channel across the plasma membrane.			
CC	-!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-bungarotoxin. The structure is probably pentameric (By similarity).			
CC	-!- SUBCELLULAR LOCATION: Integral membrane protein.			
CC	-!- SIMILARITY: Belongs to the ligand-gated ionic channel family.			
CC	-----			
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).			
CC	-----			
DR	EMBL; L37663; AAC42053.1; -			
DR	PIR; A57175; A57175.			
DR	MGI; MG1:99779; Chrna7.			
DR	InterPro; IPR006029; Neu_chan memb.			
DR	InterPro; IPR006202; Neu_chan LBD.			
DR	InterPro; IPR006201; Neu_chan.			
DR	Pfam; PF02931; Neu_chan LBD; 1.			
DR	Pfam; PF02932; Neu_chan memb; 1.			
DR	PRINTS; PR00252; NRIONCHANNEL.			
DR	TIGRFAMS; TIGR00860; LIC; 1.			
DR	PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.			
KW	Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;			
KW	Transmembrane; Multigene family.			
FT	SIGNAL 1 22			
FT	CHAIN 23 502			
FT	BY SIMILARITY.			
FT	NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,			
FT	ALPHA-7 CHAIN.			
FT	EXTRACELLULAR (POTENTIAL).			
FT	DOMAIN 23 230			

FT TRANSMEM 231 255 POTENTIAL.
FT TRANSMEM 262 280 POTENTIAL.
FT TRANSMEM 296 317 POTENTIAL.
FT DOMAIN 318 469 CYTOPLASMIC (POTENTIAL).
FT TRANSMEM 470 490 POTENTIAL.
FT FT DISULFID 150 164 BY SIMILARITY.
FT FT DISULFID 212 213 ASSOCIATED WITH RECEPTOR ACTIVATION
(BY SIMILARITY).
FT CARBOHYD 46 46 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL).
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ SEQUENCE 502 AA; 56631 MW; C9312E5226D120E3 CRC64;
Query Match 29.28; Score 1182.5; DB 1; Length 502;
Best Local Similarity 45.8; Pred. No. 5.1e-75;
Matches 239; Conservative 80; Mismatches 118; Indels 85; Gaps 10;
QY 297 IYLNLSA---KVLGAGYHKKLLHLDLPYNTLPRVNESDPLQLSFLGLTLMQIIDVDE 353
Db 8 IWLALAAALHVSLOGEFQRLYKELVKNYNPLRPVANDSQPLTVYFVLSLLQIMDVDE 67
QY 354 KNOLLVTNWLKLEWDMNLRWNTSDYGVKQRLRIPPHRIWKPDVLMYNSADEGFDGTQY 413
Db 68 KNOVLATTNIWLQMSWTBHYLQWNNSEYPGVKNVRFDPGQIWKPDILLYNSADERFDATEH 127
QY 414 TNNVVRNNGSLVYPPGIFKSTCKIDITWPFDDORCEMKEGSWTYDGFOLDILODETG 473
Db 128 TNNLVNASGHCQYLPPIGIFKSSCYIDVRWPFDDVQCKIKFGSWSYGWSLIDLQMBE--- 184
QY 474 GDTSSVYNGEWELLGVPGRKNEYVNCPEPYDITFAIIRRTTYFFNLIIPCVLI 533
Db 185 ADISSYTPNGEWDLMGIPGRKNEYFCCKEPEYDVTYVYTMRRRTLYGLNLLIPCVLI 244
QY 534 ASWALLGFTLPPDGEKLSLGVITLLSLTVPLNMVAETMPATSDAVPI----- 581
Db 245 SALALLVFLPADSGEKISLITVLLSLTVPLMVAETMPATSDVPLIAQYFASMTIIV 304
QY 582 -----WIRIVFLCWLPLWLRSPG---RELILFEFTTP 612
Db 305 GLSVVVTVIVLYHHHPDGGKMPKWTIRILLNCWFLRMKRGEGEDKVRACQHKPR- 363
QY 613 CSUTSRRKHQILSDVELKERSKSLANVLIDDDFRHNCRPMTPGGTFLPHNPAFYRTV 672
Db 364 CS-----LASVELSAGAG-----PRTSGNGLLY--IGRGL 392
QY 673 YGQ-----GDGSI-GPIGSTRMPDVAHTFCIKSTSEYELGLILKEIRFTDQLRDXDE 726
Db 393 EGMHCAPTSDSGVVCGRACSPTHDEHLMHGTHPSDGDPLAKILEEVRYIANFRQDCE 452
QY 727 CNDTANDWKAAMVVDRLCLIFTFMFAILATIAVLLSAPHII 768
Db 453 SEVICSEWKFACVVDRLCLMAFSVFTICTIGILMSAPNEV 494
RESULT 2
ACH7 CHICK
ID -ACH7 CHICK STANDARD; PRT; 502 AA.
AC P22770;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-7 chain precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=91097796; PubMed=1702646;
RA Couturier S., Bertrand D., Matter J.M., Hernandez M.C., Bertrand S.,
RA Millar N., Valera S., Barkas T., Ballivet M.;

"A neuronal nicotinic acetylcholine receptor subunit (alpha 7) is
developmentally regulated and forms a homo-oligomeric channel blocked
by alpha-BTX." Neuron 5:847-856(1990).
[2]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=90315158; PubMed=2369519;
RA Schoeffer R., Conroy W.G., Whiting P., Gore M., Lindstroem J.;
RT "Brain alpha-bungarotoxin binding protein cDNAs and MABs reveal
subtypes of this branch of the ligand-gated ion channel gene
superfamily." Neuron 5:35-48(1990).
[3]
RP SEQUENCE OF 1-18 FROM N.A.
RC STRAIN=White leghorn; TISSUE=Erythrocyte;
RX MEDLINE=93049204; PubMed=1425587;
RA Matter-Sadzinski L., Hernandez M.C., Rostocil T., Ballivet M.,
RA Matter J.M.;
RT "Neuronal specificity of the alpha 7 nicotinic acetylcholine receptor
promoter develops during morphogenesis of the central nervous
system." EMBO J. 11:4529-4538(1992).
[4]
RP SEQUENCE OF 24-47.
RC TISSUE=Brain;
RX MEDLINE=85270494; PubMed=3860855;
RA Conti-Tronconi B.M., Dunn S.M.J., Barnard E.A., Dolly J.O., Lai F.A.,
RA Ray N., Rafferty M.A.;
RT "Brain and muscle nicotinic acetylcholine receptors are different but
homologous proteins." Proc. Natl. Acad. Sci. U.S.A. 82:5208-5212(1985).
[5]
RP MUTAGENESIS OF LEU-270.
RX MEDLINE=92049732; PubMed=1719423;
RA Recah F., Bertrand D., Galzi J.-L., Devillers-Thiery A., Mulle C.,
RA Hussy N., Bertrand S., Ballivet M., Changeux J.-P.;
RT "Mutations in the channel domain alter desensitization of a neuronal
nicotinic receptor." Nature 353:846-849(1991).
[6]
RP MUTAGENESIS TO CONVERT ION SELECTIVITY FROM CATIONIC TO ANIONIC.
RX MEDLINE=93024917; PubMed=1383829;
RA Galzi J.-L., Devillers-Thiery A., Hussy N., Bertrand S.,
RA Changeux J.-P., Bertrand D.;
RT "Mutations in the channel domain of a neuronal nicotinic receptor
convert ion selectivity from cationic to anionic." Nature 359:500-505(1992).
CC -!- FUNCTION: After binding acetylcholine, the ACHR responds by an
extensive change in conformation that affects all subunits and
leads to opening of an ion-conducting channel across the plasma
membrane.
CC -!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-
bungarotoxin. The structure is probably pentameric (By
similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- DEVELOPMENTAL STAGE: Alpha-7 transcripts transiently accumulate
in the developing optic tectum between E5 and E16.
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
or send an email to license@sib-sib.ch).

CC EMBL; X52295; CAA36543.1; -;
CC EMBL; X68246; CAA48317.1; -;
CC EMBL; X68586; CAA48576.1; -;
CC PIR; JN0113; JN0113.
CC PDB; 1KC4; 17-APR-02.

```
DR PDB; 1KL8; 17-APR-02.
DR InterPro; IPR006029; Neu_chan_memb.
DR InterPro; IPR006202; Neu_chan_LBD.
DR InterPro; IPR006201; Neu_chan_LBD.
DR Pfam; PF02931; Neu_chan_LBD; 1.
DR Pfam; PF02932; Neu_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family; 3D-structure.
FT SIGNAL 1 23
FT CHAIN 24 502
FT DOMAIN 24 230
FT TRANSMEM 231 255
FT TRANSMEM 262 280
FT TRANSMEM 296 317
FT DOMAIN 318 469
FT TRANSMEM 470 490
FT DISULFID 150 164
FT DISULFID 212 213
FT ASSOCIATED WITH RECEPTOR ACTIVATION
FT (BY SIMILARITY).
FT N-LINKED (GLNAC. . .) (POTENTIAL).
FT N-LINKED (GLNAC. . .) (POTENTIAL).
FT L-S, T; SUPPRESSES INHIBITION BY THE
FT OPEN-CHANNEL BLOCKER OX-222.
FT OR -> ET (IN REF. 3).
FT CONFLICT 26 27
FT SEQUENCE 502 AA; 56946 MW; 572325D4309AD2FD CRC64;
Query Match 29.2%; Score 1179.5; DB 1; Length 502;
Best Local Similarity 45.7%; Pred. No. 8.2e-75;
Matches 237; Conservative 80; Mismatches 127; Indels 75; Gaps 9;
QY 295 LTIYLNLSA---KVLGAGYHEKELLDLPYNTLPRVNLNEDPQLSLGFLTMQIDV 351
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
6 LMLWLLAAGLVRESLQGEFQKLYKXLLKNYPLRPVANDSQPLTVFTLSLMQMDV 65
QY 352 DEKNQLLVTNWLKLEWMDNMLWNTSDYGVKDLRIPPHRIWKPDLVLMYNSADEFGDT 411
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
66 DEKNQVLTNMLQMYTHTYLOMNVSEYGVKNRVFPDGLIWKPDILLYNSADERFDT 125
QY 412 YQTNVVRNNGSLYVPPGFKSTCKDITWFPDQRCMKFGSWTYDGFQDLQLODE 471
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
126 FHTNVLNMSGHCQYLLPPGIFKSSCIDVWFPPDVQKNLFGSWTYGWSLDLQMQE- 184
QY 472 TGGDISYVINGEWELLGVPGKNEIYNCPPYIDITPAITIRRTLYFFNLIPCV 531
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
185 --ADISGYISNGEWDLVGIPGKTESFYECCKEPYDITFTVMRRRTLYYGLNLIPCV 242
QY 532 LIASMLLGFLLPDSGEKLSIGVTLLSLTVPLNVAETMPATSDAVPL----- 581
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
243 LISALALLVFLPADSGEKISLIGITVLLSLTVPMLLVAETMPATSDVPLIAQYFASTMI 302
QY 582 -----WIRIVFLWMLPWLIRMSRRCRGLLLEFPPTPCS 614
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
303 IVGLSVVTVIVLYQHHPDDGKMPKTRVILLNWCAMFLRMKRPG-----EDKVRPAC 357
QY 615 DTSSERKHQILSDVELKERSKSL--ANVLIDIDDDFRH-----NCRWPTPGTLPHNPAFY 669
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
358 QHKQRCS--LSSMENVTSGQCSNGNMLYI--GFRGLDGVHCTTTDSGVI----- 406
QY 670 RTVYGGDDGSGIPGISTMPDVAHTHTCKISTEYELGLIKBIFITDQLRKDECDND 729
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
407 -----CGRMTCSPTTEENLHSGHSPSGDPDLAKILLEEVYIANRFRDQDEEA 455
QY 730 IANDWKAFAWVDRCLLIIFTFAILLATIAVLLSAPHII 768
Db : : : : : : : : : : : : : : : : : : : : : : : : : : : :
456 ICNEWKFAASVVDRLCLMAFSVFTICTIGILMSAPNFV 494
RESULT 3
```

```
FT CARBOHYD 46 46 (BY SIMILARITY)
FT CARBOHYD 90 90 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 133 133 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CARBOHYD 447 447 N-LINKED (GLCNAC. . .) (POTENTIAL)
FT CONFLICT 447 447 N -> F (IN REF. 2)
FT CONFLICT 469 469 P -> R (IN REF. 2)
SQ SEQUENCE 502 AA; 56410 MW; 00996E74EC7B9A56 CRC64;

Query Match 29.08; Score 1173.5; DB 1; Length 502;
Best Local Similarity 45.68; Pred. No. 2.1e-74;
Matches 238; Conservative 80; Mismatches 119; Indels 85; Gaps 10;

QY 297 IYLNLSA---KVLGAGYHEKRLHDLDPYNTLPRVNLNEDPLQLSFLGLTLMQIIVDVE 353
DQ 8 IWLALAAALHVSLOGEFORRLYKELVKNYNFLERPVANDSQPLTVFSLSLQIMDVE 67
QY 354 KNQLLVTVNLKLEWDMNLRWNTSDYGVKVDLRIPPHRIWKPVLVLYNLSADEGFDGTYQ 413
DQ 68 KNQVLTTNWLQSWTQHYLQWNNSEYPGVKNVRFDPGQIQWKPDPILLYNLSADERFDATEH 127
QY 414 TNVVVRNNGSLVYVPGIEFKSTCKIDITWPEDDORCEMKGSTWYDGFOLDLQLODETG 473
DQ 128 TNLVNVNASHGHCQYLPFGIFKSSCIDVRWFPFDVQCKLXPGSWSGWSDLDQMGE--- 184
QY 474 GDISSYVLNGEWELLGVPGKRNTEYVNCPEPYDITFAIIRRTLYFFNLIPCVLI 533
DQ 185 ADISSYTPNGEWDLMGIPGKNEKFECKEPEYDVTYVTRRTLYYGLNLLIPCVLI 244
QY 534 ASWALLGFTLPDSEKLSLGVITLLSLTVPLNVAETMPSATSDAVPL----- 581
DQ 245 SALALLVFLPADSGEKISLITVLLSLTVFLLVAETMPSATSDVPLIAGYFASWTIIV 304
QY 582 -----WIRIVLWLPWILMSRPG-----RPLILEFTTP 612
DQ 305 GLSVVVTVIVLYRHHDPDGKMPKWTIRILLNCAWFLRKRGEDKVPACQHKPRR- 363
QY 613 CSDTSSEKXHQILSDVELKERSSKLLANVLIDDDFFHRCRPTPGTLPHPNPAFRTV 672
DQ 364 CS-----LASVELSAGAG-----PPTSNGNLLY--IGRGL 392
QY 673 YGO-----GDGSI-GPIGSTRMPDVAHTHTCKSTSEYELGLILKEIRFITDQLRKDD 726
DQ 393 EGMHCAPTSPGSGVVCGLACSPTHDEHLMHGAHPSDGDPLAKILEEVRYIANRRCQDE 452
QY 727 CNDTANDWKAAMVVDRLCLIFTFMEALLATIAVLLSAPHII 768
DQ 453 SEVICSEWKFACVVDPLCLMAFVSFTICTIGILMSAPNFV 494

RESULT 4
ACH7 HUMAN
ID ACH7_HUMAN STANDARD; PRT; 502 AA.
AC P36544; Q15826; Q96KH2; Q99555; Q9BXH0;
DT 01-JUN-1994 (Rel. 29, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-7 chain precursor.
GN CHRNA7 OR NACHRA7.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI TaxID=9606;
[1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=94195283; PubMed=8145738;
RA Peng X., Katz M., Gerzanich V., Anand R., Lindstrom J.;
RT "Human alpha 7 acetylcholine receptor: cloning of the alpha 7 subunit
RT from the SH-SY5Y cell line and determination of pharmacological
RT properties of native receptors and functional alpha 7 homomers
RT expressed in Xenopus oocytes."
RL Mol. Pharmacol. 45:546-554 (1994).
RN [2]

SEQUENCE FROM N.A.
TISSUE=Hippocampus;
Logel J., Drebing C., Barnhart M., Antle C., Leonard S.;
Submitted (DEC-1995) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RX MEDLINE=97062879; PubMed=8906617;
RA Elliott K.J., Ellis S.B., Berckhan K.J., Urrutia A.,
RA Chavez-Noriega L.E., Johnson E.C., Velicelebi G., Harpold M.M.;
RT "Comparative structure of human neuronal alpha 2-alpha 7 and beta
RT 2-beta 4 nicotinic acetylcholine receptor subunits and functional
RT expression of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and
RT beta 4 subunits."
RA J. Mol. Neurosci. 7:217-228 (1996).
[4]
RP SEQUENCE FROM N.A.
RX MEDLINE=97162233; PubMed=9009220;
RA Groot Kormelink P.J., Luyten W.H.M.L.;
RT "Cloning and sequence of full-length cDNAs encoding the human neuronal
RT nicotinic acetylcholine receptor (nAChR) subunits beta3 and beta4 and
RT expression of seven nAChR subunits in the human neuroblastoma cell
RT line SH-SY5Y and/or IMR-32";
RA FEBS Lett. 400:309-314 (1997).
[5]
RP REVISIONS.
RA Groot Kormelink P.J., Luyten W.H.M.L.;
RA Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
[6]
RP SEQUENCE FROM N.A.
TISSUE=Epidermal keratinocytes;
RA Arredondo J., Grando S.A.;
RT "Cloning cholinergic receptors in human keratinocytes."
RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
[7]
RP SEQUENCE OF 17-502 FROM N.A.
TISSUE=Brain;
RA Doucette-Stamm L., Monteggia L.M., Donnelly-Roberts D., Wang M.T.,
RA Lee J., Tian J., Giordano T.;
RT "Cloning and sequence of the human a7 nicotinic acetylcholine
RT receptor."
RA Drug Dev. Res. 30:252-256 (1993).
[8]
RP SEQUENCE OF 24-502 FROM N.A.
TISSUE=Retina;
RA Chini B., Raimondi E., Elgoyhen A.B., Moralli D., Balzaretto M.,
RA Heinemann S.F.;
RT "Molecular cloning and chromosomal localization of the human alpha 7-
RT nicotinic receptor subunit gene (CHRNA7)."
RA Genomics 19:379-381 (1994).
[9]
RP SEQUENCE OF 118-129 FROM N.A.
RX MEDLINE=21818878; PubMed=11829490;
RA Riley B., Williamson M., Collier D., Wilkie H., Makoff A.;
RT "A 3-Mb map of a large segmental duplication overlapping the alpha7-
RT nicotinic acetylcholine receptor gene (CHRNA7) at human 15q13-q14."
RA Genomics 79:197-209 (2002).
[10]
RP MASS SPECTROMETRY.
TISSUE=Breast cancer;
RX MEDLINE=21829512; PubMed=11840567;
RA Harris R.A., Yang A., Stein R.C., Lucy K., Brusten L., Herath A.,
RA Parekh R., Waterfield M.D., O'Hare M.J., Neville M.A., Page M.J.,
RA Zvelebil M.J.;
RT "Cluster analysis of an extensive human breast cancer cell line
RT protein expression map database."
RA Proteomics 2:212-223 (2002).
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
CC -!- SUBUNIT: Forms a homo-oligomeric channel blocked by alpha-
CC bungarotoxin. The structure is probably pentameric (B)
```

This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).

```

RESULT 9
ACHI_SCHGR
ID ACHI_SCHGR STANDARD; PRT; 557 AA.
P23414;
AC DT 01-NOV-1991 (Rel. 20, Created)
AT DT 01-NOV-1991 (Rel. 20, last sequence update)
DT DT 28-FEB-2003 (Rel. 41, last annotation update)
DE Acetylcholine receptor protein, alpha-L1 chain precursor.
OS Schistocerca gregaria (Desert locust).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Orthopteroidea; Orthoptera; Caelifera; Acridomorpha;
OC Acridoidea; Acrididae; Cyrtacanthacridinae; Schistocerca.
OC NCBI_TaxID=7010;
RX [1]
RN SEQUENCE FROM N.A.
RP MEDLINE=91092263; PubMed=1702381;
RRA Marshall J., Buckingham S.D., Shingai R., Lunt G.G., Goosey M.W.,
RRA Darlison M.G., Sattelle D.B., Barnard E.A.;
RRT "Sequence and functional expression of a single alpha subunit of an
RRT insect nicotinic acetylcholine receptor.";
RRL EMBO J. 9:4391-4398(1990).
RRC
RCC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
RCC extensive change in conformation that affects all subunits and
RCC leads to opening of an ion-conducting channel across the plasma
RCC membrane.
RCC -!- SUBCELLULAR LOCATION: Integral membrane protein.
RCC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
RCC
RCC This SWISS-PROT entry is copyright. It is produced through a collaboration
RCC between the Swiss Institute of Bioinformatics and the EMBL outstation
RCC the European Bioinformatics Institute. There are no restrictions on its
RCC use by non-profit institutions as long as its content is in no way
RCC modified and this statement is not removed. Usage by and for commercial
RCC entities requires a license agreement (See http://www.isb-sib.ch/announce/
RCC or send an email to license@isb-sib.ch).
RCC
RCC EMBL; X55439; CA339081.1; -.
RCC PIR; S12359; S12359.
RCC InterPro; IPR006029; Neu channel memb.
RCC

```

DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 28-FEB-2003 (Rel. 41, Last annotation update)
 DE Neuronal acetylcholine receptor protein, alpha-2 chain precursor.
 GN CHRNA2.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP TISSUE=Hypothalamus;
 RC MEDLINE=97062879; PubMed=8906617;
 RX Elliott K.J., Ellis S.B., Bercikhan K.J., Urrutia A.,
 RA Chavez-Noriega L.E., Johnson B.C., Velicelebi G., Harpold M.M.;
 RT "Comparative structure of human neuronal alpha 2-alpha 7 and beta
 2-beta 4 nicotinic acetylcholine receptor subunits and functional
 expression of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and
 beta 4 subunits.";
 RT J. Mol. Neurosci. 7:217-228(1996).
 RL [2]
 RN SEQUENCE FROM N.A.
 RP Groot Kormelink P.J.;
 RA Submitted (JAN-1998) to the EMBL/GenBank/DBJ databases.
 RL [3]
 RN SEQUENCE FROM N.A.
 RA Blechschmidt K., Rosenthal A.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
 CC extensive change in conformation that affects all subunits and
 CC leads to opening of an ion-conducting channel across the plasma
 CC membrane.
 CC -!- SUBUNIT: Neuronal AChR seems to be composed of two different types
 CC of subunits: alpha and non-alpha (beta). Alpha-2 subunit can be
 CC combined to beta-2 or beta-4 to give rise to functional receptors.
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
 CC
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; U62431; AAB40109.1; -;
 DR EMBL; Y16281; CAA76154.1; -;
 DR EMBL; AF311103; -; NOT ANNOTATED_CDS.
 DR Genew; HGNC:1956; CHRNA2.
 DR MIM; 118502; -;
 DR GO; GO:0005892; C:nicotinic acetylcholine-gated receptor-chan. . . ; TAS.
 DR GO; GO:0013464; F:fractylcholine receptor activity; TAS.
 DR GO; GO:0004889; F:nicotinic acetylcholine-activated cation-se. . . ; TAS.
 DR GO; GO:0007165; P:signal transduction; TAS.
 DR GO; GO:0007268; P:synaptic transmission; TAS.
 DR InterPro; IPR006029; Neu_chan_memb.
 DR InterPro; IPR006202; Neu_chan_LBD.
 DR InterPro; IPR006201; Neu_chan_LBD.
 DR Pfam; PF02931; Neur_chan_LBD; 1.
 DR Pfam; PF02932; Neur_chan_memb; 1.
 DR PRINTS; PR00252; NRIONCHANNEL.
 DR TIGRFAMs; TIGR00860; LIC; 1.
 DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
 KW Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane; Multigene family.
 FT SIGNAL 1 26 POTENTIAL.
 FT CHAIN 27 529 NEURONAL ACETYLCHOLINE RECEPTOR PROTEIN,
 FT ALPHA-2 CHAIN.
 FT DOMAIN 27 264 EXTRACELLULAR.
 FT TRANSMEM 265 289 POTENTIAL.
 FT TRANSMEM 297 315 POTENTIAL.
 FT TRANSMEM 331 352 POTENTIAL.
 FT DOMAIN 353 502 CYTOPLASMIC.

FT TRANSMEM 503 521 POTENTIAL.
 FT DISULFID 183 197 BY SIMILARITY.
 FT DISULFID 247 248 ASSOCIATED WITH RECEPTOR ACTIVATION
 FT (BY SIMILARITY).
 FT CARBOHYD 79 79 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 129 129 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CARBOHYD 235 235 N-LINKED (GLCNAC. . .) (POTENTIAL).
 FT CONFLICT 125 125 A -> T (IN REF. 3).
 SQ SEQUENCE 529 AA; 59735 MW; 7F512B06CCD9AAFD CRC64;
 Query Match 21.0%; Score 848.5; DB 1; Length 529;
 Best Local Similarity 36.7%; Pred. No. 1e-51;
 Matches 183; Conservative 87; Mismatches 154; Indels 75; Gaps 10;
 QY 309 GYH---EKRLHLLDLPYNTLERPVLESPLQLSPGLTLMQIIDVDEKQNLTVNVLWK 365
 Db 53 GSHTEDEDLFKHLFRGYNRWARPVNTSDVIVRFGLSIAQLIDVDEKQMTTNNVLK 112
 QY 366 LEWNDMLRWNTSDYGVKDLRIPIPHRWKPDVLMYNSADEGFDGTQTNVVVRNNGSCL 425
 Db 113 QEWSDYKLRWNPADFGNITSLRVSEMIWIPDIVLYNNADGEFAVTHMTKAHLFSTGVH 172
 QY 426 YVPGIEPKSTCKIDITWPPDDQCEMKFGSWTYDGFOLDLQLODETGDISSVYVNGEW 485
 Db 173 WPPAIYKSSCSIDVTFPPDQCKMKFGSWTYDKAKIDLEQMEQT-VDLKDYWESGEW 231
 QY 486 ELLGVGKRNEIYVNCPEYIDITFAIIRRTLYYFFNLIIPCVLIASVALLGFTLPP 545
 Db 232 AIVNATGYNSKYDCCAEIYPDVYAFVIRRLPLFTVNIIPCLISCLTVLVFYLPS 291
 QY 546 DSGKLSLGVILLISLTVFLNMVAETMPATSDAVPL----- 581
 Db 292 DCGEKITLCISVLSLTVFLLLITELIPSTSLVPLIGEYLLFTMTFTLSIVITVFLVN 351
 QY 582 -----WIRVFLCWLPMILRMSPGRPLILEFTPTTCSDTSSRKHQILSDV 628
 Db 352 VHRSPSTHTPHWVRGALLCGVPRWLMNRPPEVEL---CHPLRLKLSPYHWSNV 408
 QY 629 ELKERSKSLANVLIDDDFRHNCR-PMTPG-GTLPHPNAPYRTVYGGDGSIGPTGS 686
 Db 409 DAEERE-----VVVEEDRWACAGHVAPSVGTL-----CSGHLHSGASGPKAE 452
 QY 687 TRMPDA---VTHHTCIKSTVELGLILKEIRFIDQLRKDKDCNDIANDKFAAMVVD 743
 Db 453 ALLQGEILLSPH-----MOKALEGVHYIADHLRSEDADSSVKEDWKYVAVIDR 502
 QY 744 LCLLIIFTMFAILLATIAVLL 762
 Db 503 IFLWLFIVCFGLTIGLFL 521

RESULT 11

ACH3 HUMAN
 ID ACH3 HUMAN STANDARD; PRT: 503 AA.
 AC P32977; Q15823; Q96RH3; Q99553; Q9BQ93;
 DT 01-OCT-1993 (Rel. 27, Created)
 DT 01-NOV-1997 (Rel. 35, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Neuronal acetylcholine receptor protein, alpha-3 chain precursor.
 GN CHRNA3 OR NACHRA3
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RN SEQUENCE FROM N.A.
 RP MEDLINE=90245296; PubMed=2336208;
 RX Fornasari D., Chini B., Tarroni P., Clementi F.;
 RA "Molecular cloning of human neuronal nicotinic receptor alpha
 RT 3-subunit";
 RL Neurosci. Lett. 111:351-356(1990).
 RN [2]
 RP SEQUENCE FROM N.A. (ISOFORM 1).

RC TISSUE=Thymus;
 RX MEDLINE=911114756; PubMed=1989896;
 RA Mihovilovic M., Roses A.D.;
 RT "Expression of mRNAs in human thymus coding for the alpha 3 subunit
 of a neuronal acetylcholine receptor.";
 RL Exp. Neurol. 111:175-180(1991).
 RN [3]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=97062879; PubMed=8906617;
 RA Elliott K.J., Ellis S.B., Berckhan K.J., Urrutia A.,
 RA Chavez-Noriega L.E., Johnson E.C., Velicelebi G., Harpold M.M.;
 RT "Comparative structure of human neuronal alpha 2-alpha 7 and beta
 2-beta 4 nicotinic acetylcholine receptor subunits and functional
 expression of the alpha 2, alpha 3, alpha 4, alpha 7, beta 2, and
 beta 4 subunits.";
 RL J. Mol. Neurosci. 7:217-228(1996).
 RN [4]
 RP SEQUENCE FROM N.A. (ISOFORM 1).
 RX MEDLINE=97162233; PubMed=9009220;
 RA Groot Kormelink P.J., Luyten W.H.M.L.;
 RT "Cloning and sequence of full-length cDNAs encoding the human neuronal
 nicotinic acetylcholine receptor (nAChR) subunits beta3 and beta4 and
 expression of seven nAChR subunits in the human neuroblastoma cell
 line SH-SY5Y and/or IMR-32.";
 RL FEBS Lett. 400:309-314(1997).
 RN [5]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=99118870; PubMed=9921897;
 RA Rempel N., Heyers S., Engels H., Slegers E., Steinlein O.K.;
 RT "The structures of the human neuronal nicotinic acetylcholine receptor
 beta2- and alpha3-subunit genes (CHRNA2 and CHRNA3).";
 RL Hum. Genet. 103:645-653(1998).
 RN [6]
 RP SEQUENCE FROM N.A. AND VARIANT LEU-21 INS.
 RX MEDLINE=21342809; PubMed=11450844;
 RA Lev-Lehman E., Bercovich D., Xu W., Stockton D.W., Beaudet A.L.;
 RT "Characterization of the human beta4 nAChR gene and polymorphisms in
 CHRNA3 and CHRNA4.";
 RL J. Hum. Genet. 46:362-366(2001).
 RN [7]
 RP SEQUENCE FROM N.A. (ISOFORM 2).
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins P.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uesdin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Villalón D.K., Muzny K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Fahy J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickinson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smallus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [8]
 RP SEQUENCE OF 30-503 FROM N.A.
 RX TISSUE=Brain;
 RA Anand R., Lindstrom J.;
 RT Submitted (JUN-1990) to the EMBL/GenBank/DBJ databases.
 RN [9]
 RP SEQUENCE OF 6-493 FROM N.A.
 RX TISSUE=Epidermal keratinocytes;
 RA Arredondo J., Grando S.A.;
 RT "Cloning cholinergic receptors in human keratinocytes.";

Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
 -!- FUNCTION: After binding acetylcholine, the AChR responds by an
 extensive change in conformation that affects all subunits and
 leads to opening of an ion-conducting channel across the plasma
 membrane.
 -!- SUBUNIT: Neuronal AChR seems to be composed of two different type
 of subunits: alpha and non-alpha (beta).
 -!- SUBCELLULAR LOCATION: Integral membrane protein.
 -!- ALTERNATIVE PRODUCTS:
 Event=Alternative splicing; Named isoforms=2;
 Name=1;
 IsoId=P32297-1; Sequence=Displayed;
 Name=2;
 IsoId=P32297-2; Sequence=VSP_000073;
 Note=No experimental confirmation available;
 -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.

 This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL Outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).

 EMBL; M86383; AAC84176.1; -
 EMBL; M37981; AAA59942.1; -
 EMBL; U62432; AAB40110.1; -
 EMBL; Y08418; CAA69695.1; -
 EMBL; AJ007783; CAA07682.1; -
 EMBL; AJ007784; CAA07682.1; JOINED.
 EMBL; AJ007785; CAA07682.1; JOINED.
 EMBL; AJ007786; CAA07682.1; JOINED.
 EMBL; AJ007787; CAA07682.1; JOINED.
 EMBL; BC001642; AAH01642.1; -
 EMBL; BC002996; AAH02996.1; -
 EMBL; BC000513; AAH00513.1; -
 EMBL; AF385584; AAK68110.1; -
 EMBL; X53559; CAA37625.1; -
 PIR; A37040; A37040.
 PIR; A53956; A53956.
 Genew: HGNC:1957; CHRNA3.
 MIM; 118503; -
 GO; GO:0005892; Nicotinic acetylcholine-gated receptor-chan. . . ; TAS.
 GO; GO:0015464; Facetylcholine receptor activity; TAS.
 GO; GO:0004889; F:nicotinic acetylcholine-activated cation-se. . . ; TAS.
 GO; GO:0005215; F:transporter activity; TAS.
 GO; GO:0007165; P:signal transduction; TAS.
 GO; GO:0006810; P:transport; TAS.
 InterPro; IPR006029; Neu_chan_memb.
 InterPro; IPR006202; Neur_chan_LBD.
 InterPro; IPR006201; Neur_chan.
 Pfam; PF02931; Neur_chan_LBD; 1.
 Pfam; PF02932; Neur_chan_memb; 1.
 PRINTS; PR00252; NRIONCHANNEL.
 TIGRfams; TIGR00860; LIC; 1.
 PROSITE; PS00236; NEUOTR_ION_CHANNEL; 1.
 Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 Transmembrane; Multigene family; Alternative splicing; Polymorphism.
 SIGNAL 1 29
 CHAIN 30 503
 DOMAIN 30 238
 TRANSMEM 239 263
 TRANSMEM 271 289
 TRANSMEM 305 326
 DOMAIN 327 475
 TRANSMEM 476 495
 DISULFID 157 171
 DISULFID 221 222
 CARBOHYD 53 53
 CARBOHYD 170 170

CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
 CC -!- TISSUE SPECIFICITY: CNS in embryos.
 CC -!- DEVELOPMENTAL STAGE: Late embryonic and late pupal stages.
 CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; X52274; CAA36517.1; -;
 DR EMBL; X53583; CAA37652.1; -;
 DR EMBL; AE003748; AAF56303.1; -;
 DR EMBL; AY058446; AAL13675.1; -;
 DR PIR; S11679; ACFFA2.
 DR FlyBase; FBgn0000039; nAcr-alpha-96Ab.
 DR InterPro; IPR006029; Neur channel memb.
 DR InterPro; IPR006202; Neur chan LBD.
 DR InterPro; IPR006201; Neur channel.
 DR Pfam; PF02931; Neur_chan_LBD; 1.
 DR Pfam; PF02932; Neur_chan_memb; 1.
 DR PRINTS; PR00252; NRIONCHANNEL.
 DR TIGRFAMs; TIGR00860; LIC; 1.
 DR PROSITE; PS00236; NEUROTROPHIC ION CHANNEL; 1.
 DR Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
 KW Transmembrane; Multigene family.
 FT SIGNAL 1 21 PROBABLE.
 FT CHAIN 22 576 ACETYLCHOLINE RECEPTOR PROTEIN, ALPHA-
 FT LIKE CHAIN 2.
 FT DOMAIN 22 261 EXTRACELLULAR (POTENTIAL).
 FT TRANSMEM 262 285 POTENTIAL.
 FT TRANSMEM 293 311 POTENTIAL.
 FT TRANSMEM 327 346 POTENTIAL.
 FT DOMAIN 347 526 CYTOPLASMIC (POTENTIAL).
 FT TRANSMEM 527 545 POTENTIAL.
 FT DISULFID 169 183 BY SIMILARITY.
 FT DISULFID 243 244 ASSOCIATED WITH RECEPTOR ACTIVATION
 FT (BY SIMILARITY).
 FT CARBOHYD 65 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 254 N-LINKED (GLCNAC. .) (POTENTIAL).
 FT CARBOHYD 570 N-LINKED (GLCNAC. .) (POTENTIAL).
 SQ SEQUENCE 576 AA; 65506 MW; 97D6A46CADCF42F CRC64;
 Query Match 20.9%; Score 843.5; DB 1; Length 576;
 Best Local Similarity 34.4%; Pred. No. 2.6e-51;
 Matches 193; Conservative 87; Mismatches 170; Indels 111; Gaps 11;
 QY 290 HSW-----IFLLIVNLSAKVCLAGYHEKRLHDLDPYNTLRPPVLNESDPLQSLGT 344
 Db 18 HIWCKPGLCLLVLLLCCTVQANPPAKLYDLDLSNRLIRPVSNNTDTLVKLGLR 77
 QY 345 LMQIIDVDEKQLLVNVLKLEWMDNMLRWNTSDYGVKDLRTPPHRWKPVLMVNSA 404
 Db 78 LSQLIDLNLKQILITNVWLEHEWQDHKFDSEYGVGVTELYVPSHILPDLVLYNA 137
 QY 405 DEGFDTGYQTNVVRNNGSCLYVPPGLFKSTCKIDITWFPDQRCMKFGSWTYDGFOL 464
 Db 138 DGYVVTWTKAILHYTGKVVWTPPAIFKSKCEIDVRYFPDQTCFMKFGSWTYDQDI 197
 QY 465 DL-----QLQD-----ETGGDISSVVLNGEWELGVGPKNEIYVNCPEYIDITPAIII 515
 Db 198 DLKHSQNDKXNKEIGIDREYVPSVEMDIIGVPAERHEKYPCCAEYPDIFNITL 257
 QY 516 RRTLYFFNLLIPCVLIASMLGLFTLPDPSGKLSGLVITLLISLTVFLNMVAETMPAT 575
 Db 258 RRKTLFTVNLIIPCVGISVLSVLFVLPADSGEKLALCISLLSQTMFLLEISEIPST 317
 QY 576 SDAVPL-----WIRVFLCWLPWIRMS 598
 Db 318 STALPLLGKYLFTMLLVGLSVVITIIILNIHVKPSSTHQMREWIRSFIFKRLPKLLMR 377

QY 599 RPRGLIILEFPTTPCSDTSSERKHQILSDVELKE-----RSSKSLIANVLDDDDFRHNC 653
 Db 378 VP-----KDLLRDLAANKINYGKFSKTFEGQAL-MDE----- 409
 QY 654 RPTMPGGTLPHPAPRYTVYQGD-----DGSIGPIGS-----TRMPDA 692
 Db 410 MQMNSGGSPSLRRMQGRVGGCGNMGHVTATNFRSLGVGLGGLSTLSGYNGLPVS 469
 QY 693 VT-----HHTCIKSSTVEYELGLIKERITFDQLRKDECDNDIANDKFAAMVVDRLCL 746
 Db 470 LSGLDSDLSDAARKKYPFELEKAIHNVMFIOHMQRODEFNAEDQDQMGFVAMVMDRLFL 529
 QY 747 IIFTMFAILATIALLSAPHI 767
 Db 530 WLFMTIASLVGTFTVILGEAPSL 550
 RESULT 13
 ID ACH3_DROME STANDARD; PRT; 521 AA.
 AC P04755; Q9VZC3;
 DT 13-AUG-1987 (Rel. 05, Created)
 DT 13-AUG-1987 (Rel. 05, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Acetylcholine receptor protein, beta-like chain 1 precursor.
 GN NACR-BETA-64B OR ACRD OR ACR64B OR CG11348/CG12606.
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OK NCBI_taxid=7227;
 RN [1]
 RN SEQUENCE FROM N.A.
 RA Hermans-Borgmeyer I., Zopf D., Ryseck R.-P., Hovenmann B., Betz H.,
 RA Gundelfinger E.D.;
 RA "Primary structure of a developmentally regulated nicotinic
 RT acetylcholine receptor protein from Drosophila.";
 RL EMBO J. 5:1503-1508(1986).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=88298842; PubMed=3136037;
 RA Sawruk E., Hermans-Borgmeyer I., Betz H., Gundelfinger E.D.;
 RT "Characterization of an invertebrate nicotinic acetylcholine receptor
 RL gene: the ard gene of Drosophila melanogaster.";
 RN FEBS Lett. 235:40-46(1988).
 RN [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=88174720; PubMed=2832736;
 RA Wadsworth S.C., Rosenthal L.S., Kammermeyer K.L., Potter M.B.,
 RA Nelson D.J.;
 RT "Expression of a Drosophila melanogaster acetylcholine receptor-
 RL related gene in the central nervous system.";
 RN Mol. Cell. Biol. 8:778-785(1988).
 RN [4]
 RN SEQUENCE FROM N.A.
 RC STRAIN=Berkeley;
 RX MEDLINE=20196006; PubMed=10731132;
 RA Adams M.D., Celnikier S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
 RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
 RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
 RA Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
 RA Ballaw R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.Y., Benos P.V., Berman B.P., Bhargava M., Bolshakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,


```

RA Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodex A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D.A., Helman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Lasko P., Lei Y., Levitsky A.A., Li J.H., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Markulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong P.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.; "
RA "The genome sequence of Drosophila melanogaster."
RL Science 287:2185-2195(2000).
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- TISSUE SPECIFICITY: CNS in embryos.
CC -!- DEVELOPMENTAL STAGE: Late embryonic and late pupal stages.
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation
CC at the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
DR EMBL; X04016; CAA27641.1; -.
DR EMBL; X07956; CAA30778.1; -.
DR EMBL; X07957; CAA30778.1; JOINED.
DR EMBL; X07958; CAA30778.1; JOINED.
DR EMBL; M20316; AAA28311.1; -.
DR EMBL; AE003481; AAF47900.1; -.
DR PIR; S03012; ACFFNN.
DR FlyBase; Fgn0000038; nAChR-beta-64B.
DR InterPro; IPR006029; Neu channel memb.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006201; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUOTR ION CHANNEL; 1.
KW Receptor; Postsynaptic membrane; Ionic channel; Glycoprotein; Signal;
KW Transmembrane; Multigene family.
FT SIGNAL 1 24 POTENTIAL.
FT CHAIN 25 521 ACETYLCHOLINE RECEPTOR PROTEIN, BETA-LIKE
FT CHAIN 1.
FT FT DOMAIN 25 235
FT FT TRANSMEM 26 260
FT FT TRANSMEM 268 286
FT FT TRANSMEM 302 323
FT FT DOMAIN 324 481
FT FT TRANSMEM 482 500
FT FT DISULFID 152 166
FT FT CARBOHYD 48 48
FT VARIANT 73 73
FT CONFLICT 383 384 EL -> DV (IN REF. 3).
FT SEQUENCE 521 AA; 59901 MW; FF9BA2ABC0C3AA62 CRC64;
SQ

```

```

Query Match 20.8%; Score 840; DB 1; Length 521;
Best Local Similarity 35.6%; Pred. No. 3.9e-51;
Matches 186; Conservative 95; Mismatches 172; Indels 70; Gaps 9;

QY 291 SWIFLLIYNLSAKVCLAGYHEKRLHDLDPYNTLPRPVNLSEDPQLSLGFLTMOIID 350
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 7 SWLLCSLVLVAFSLVSASEDEERLVRDLFRGYNKLRPVQNMNTPKQVGVRRGLAFVQLIN 66

QY 351 VDEKNQLLVNMLKLEWMDMLRWNTSDYGVGVKDLRIPPHRIWKPDLVMTNSADEGPDG 410
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 67 VNEKNQVKNVWLRLVYDYLQWLDADYGGIGLVRLPDKPKWKPDVILFNNADGNVEV 126

QY 411 TYQTNVVRNNGSLYVPGIFKSTCKIDITWPFPPDQRCCKEMKFGSTYDGFQDLQL-Q 469
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 127 RYKSNVLIYPTGEVLWVPPIYQSSCTIDVTYFPDQDTCIMKFGSWTFNGDQVSLALYN 186

QY 470 DETGDDISSYVNGEWELLGVPGKNEIYNCCPEP-YIDITFAIIIRRRRLTYFFENLI 528
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 187 NKNFVDSLSDYVKSQGTWDLIEVPAVLN-VYEGDSNHPETDITFYIIIRKTLFTVNLIL 245

QY 529 PCVLIASMALIGFTLPDPSGKLSGLVTILLSLTVFLNMVAETMPATS----- 576
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 246 PTVLISFLCVLFFVLPAAEGEKVTIGISILLVFLVLLVSKILLPPTSLVLPILAKYLLF 305

QY 577 -----DAVPLMIRIVFLCWLPIWILRMSPPGRLILEPPTT 611
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 306 TFIIMTVSILVTIINWFRGPRTRHPMYVIRSGIFLHYLPAPFLFMKPRKTRLRMMEM 365

QY 612 PCSTSSERKHQILSDVELKERS-----KSLANVLDDDDDFHNC---RPMTPGGTLPH 664
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 366 PGMSMPAHPHPSYGPAPLPHISAIGKQSKQMEVMSLDDHHPNCKINRKVNSGGEI-- 423

QY 665 NPAFVRTVYGGDDSGISGPIGSTMPDAVTHHTCIKSTSEYELGLILKEIFITDQLRKD 724
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 424 -----GLGD-----GCRRESSSDILLSPEASKATE-----AVEFIAEHLRNE 462

QY 725 DECNDIANDKPAAMVDRLCLIIFTMPAILATIAVLSAPHI 767
   ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||: ||:
Db 463 DLYIOTREDWKYVAMVDRLQLYIFVITVTTAGTVGILMDAPHI 505

RESULT 14
ACH2 CHICK STANDARD; PRT; 528 AA.
AC P09480;
DT 01-MAR-1989 (Rel. 10, Created)
DT 01-MAR-1989 (Rel. 10, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Neuronal acetylcholine receptor protein, alpha-2 chain precursor.
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=88283624; PubMed=3267226;
RA Nef P., Oneysen C., Alliod C., Couturier S., Ballivet M.;
RT "Genes expressed in the brain define three distinct neuronal
RL nicotinic acetylcholine receptors."
RL EMBO J. 7:595-601(1988).
CC -!- FUNCTION: After binding acetylcholine, the AChR responds by an
CC extensive change in conformation that affects all subunits and
CC leads to opening of an ion-conducting channel across the plasma
CC membrane.
CC -!- SUBUNIT: Neuronal AChR seems to be composed of two different
CC type of subunits: alpha and non-alpha (also called beta). A
CC functional receptor seems to consist of two alpha-chains and
CC three non-alpha chains.
CC -!- SUBCELLULAR LOCATION: Integral membrane protein.
CC -!- SIMILARITY: Belongs to the ligand-gated ionic channel family.
CC

```


GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: May 7, 2004, 11:30:03 ; Search time 48.3701 Seconds

(without alignments)
5022.709 Million cell updates/sec

Title: US-09-303-232-2

Perfect score: 4043

Sequence: 1 MKNAQLKTEVDDDELWLV.....MFAILLATIAVLLSAPHIIVS 770

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Watch 100%

Listing first 45 summaries

Database :

SPTREMBL 25:*

- 1: sp_archaea:*
- 2: sp_bacteria:*
- 3: sp_fungi:*
- 4: sp_human:*
- 5: sp_invertebrate:*
- 6: sp_mammal:*
- 7: sp_mhc:*
- 8: sp_organelle:*
- 9: sp_phase:*
- 10: sp_plant:*
- 11: sp_rodent:*
- 12: sp_virus:*
- 13: sp_vertebrate:*
- 14: sp_unclassified:*
- 15: sp_virus:*
- 16: sp_bacteriaph:*
- 17: sp_archaeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Match	Query Match	Length	ID	Description
1	4014.5	99.3	807	5	Q8T7V5	Q8T7V5 drosophila
2	2613	64.6	570	5	Q9VJT9	Q9VJT9 drosophila
3	2388.5	59.1	482	5	Q8T5F5	Q8T5F5 drosophila
4	1991	49.2	542	5	Q86NN7	Q86NN7 drosophila
5	1988	49.2	545	5	Q9VW19	Q9VW19 drosophila
6	1850	45.8	496	5	Q9XZ13	Q9XZ13 drosophila
7	1625.5	40.2	494	5	Q8T7S2	Q8T7S2 drosophila
8	1622.5	40.1	494	5	Q8T7S3	Q8T7S3 drosophila
9	1621.5	40.1	494	5	Q86MN8	Q86MN8 drosophila
10	1620.5	40.1	494	5	Q8T7S1	Q8T7S1 drosophila
11	1611	39.8	523	5	Q8T7R9	Q8T7R9 drosophila
12	1609	39.8	501	5	Q9XZ14	Q9XZ14 drosophila
13	1602	39.6	509	5	Q8T7S0	Q8T7S0 drosophila
14	1502	37.2	391	5	Q9NKD2	Q9NKD2 drosophila
15	1310.5	32.4	554	5	Q9V179	Q9V179 drosophila
16	1235	30.5	525	5	Q8IPE2	Q8IPE2 drosophila

17	1197	29.6	509	13	Q800C7	Q800C7 brachydanio
18	1185.5	29.3	502	11	Q9JHD6	Q9JHD6 mus musculus
19	1174	29.0	513	13	Q7T2R9	Q7T2R9 fugu rubrip
20	1168	28.9	511	13	Q03481	Q03481 gallus gall
21	1163	28.8	555	13	Q7T2U0	Q7T2U0 fugu rubrip
22	1156.5	28.6	502	6	Q866A2	Q866A2 macaca mula
23	1147.5	28.4	486	13	Q7T2S0	Q7T2S0 fugu rubrip
24	1140	28.2	335	5	Q9NKD1	Q9NKD1 drosophila
25	1108.5	27.4	554	13	Q7T2T9	Q7T2T9 fugu rubrip
26	1072.5	26.5	474	13	Q7T2U1	Q7T2U1 fugu rubrip
27	1015	25.1	480	5	Q8I932	Q8I932 caenorhabdi
28	1009.5	25.0	461	5	P91197	P91197 caenorhabdi
29	977	24.2	554	5	O62083	O62083 caenorhabdi
30	968.5	24.0	542	5	Q18556	Q18556 caenorhabdi
31	902	22.3	537	5	Q8MUR0	Q8MUR0 apis mellif
32	897	22.2	515	5	O46133	O46133 locusta mig
33	888.5	22.0	537	5	Q9U941	Q9U941 myzus persi
34	883.5	21.9	523	5	O46128	O46128 heliothis v
35	882.5	21.8	552	5	P91765	P91765 myzus persi
36	876	21.7	568	5	Q9NFR5	Q9NFR5 drosophila
37	875	21.6	545	5	O96631	O96631 heliothis v
38	861.5	21.3	531	5	O96632	O96632 heliothis v
39	859	21.2	536	5	Q8T0Y9	Q8T0Y9 alysia cal
40	854	21.1	540	5	O46134	O46134 locusta mig
41	854	21.1	595	5	P91764	P91764 myzus persi
42	851	21.0	509	5	Q9NFX8	Q9NFX8 myzus persi
43	849	21.0	497	5	O46135	O46135 locusta mig
44	849	21.0	509	5	Q8MM21	Q8MM21 aphid gossy
45	849	21.0	529	13	Q7T2S4	Q7T2S4 fugu rubrip

ALIGNMENTS

RESULT 1

Q8T7V5	PRELIMINARY;	PRT;	807 AA.
ID	Q8T7V5		
AC	Q8T7V5;		
DT	01-JUN-2002 (TREMBLrel. 21, Created)		
DT	01-JUN-2002 (TREMBLrel. 21, Last sequence update)		
DE	01-JUN-2003 (TREMBLrel. 24, Last annotation update)		
DE	Nicotinic acetylcholine receptor Dalpha5 subunit		
GN	NACR-ALPHA-34E OR NACRALPHA-34E OR BG:DS05899.4 OR BG:DS05899.5 OR CG4498 OR CG16878 OR CG32975.		
OS	Drosophila melanogaster (Fruit fly).		
OC	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;		
OC	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;		
OC	Ephydroidea; Drosophilidae; Drosophila.		
OX	NCBI_TaxID=7227;		
RN	[1]		
RP	SEQUENCE FROM N.A.		
RX	MEDLINE=21969411; PubMed=11973307;		
RA	Grauso M., Reenan R.A., Culotto E., Sattelle D.B.;		
RT	"Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing."		
RL	Genetics 160:1519-1533(2002).		
CC	-!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).		
CC	-!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.		
DR	EMBL; AF272778; AM13390.1; -		
DR	FlyBase; FBgn0028875; NACR-alpha-34E.		
DR	GO; GO:0016021; C:intracel to membrane; IEA.		
DR	GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.		
DR	GO; GO:0005216; F:ion channel activity; IEA.		
DR	GO; GO:0030594; F:neurotransmitter receptor activity; IEA.		
DR	GO; GO:0006811; P:ion transport; IEA.		
DR	GO; GO:0007266; P:synaptic transmission; IEA.		
DR	InterPro; IPR006201; Neur_chan.		
DR	InterPro; IPR006202; Neur_chan_LBD.		
DR	InterPro; IPR006029; Neur_chan_memb.		
DR	Pfam; PF029311; Neur_chan_LBD; 1.		
DR	Pfam; PF02932; Neur_chan_memb; 1.		

DR PRINTS: PR00252; NRIONCHANNEL.
DR TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 807 AA; 91223 MW; C8B4F6B34287C8C8 CRC64;
Query Match 99.3%; Score 4014.5; DB 5; Length 807;
Best Local Similarity 95.4%; Pred. No. 0;
Matches 770; Conservative 0; Mismatches 0; Indels 37; Gaps 1;
QY 1 MKNAQLKLTVEVDDDELAVLRLAHCSSNFSSSSSTRITSSNQRINQQLTLQPRSLSTKH 60
DB 1 MKNAQLKLTVEVDDDELAVLRLAHCSSNFSSSSSTRITSSNQRINQQLTLQPRSLSTKH 60
QY 61 HSNIASQHNSSQOQEPASKDEVDVANHGSRNDQOHLQOLDSSNMLSPKTAATAAGDEA 120
DB 61 HSNIASQHNSSQOQEPASKDEVDVANHGSRNDQOHLQOLDSSNMLSPKTAATAAGDEA 120
QY 121 TQOPTNIRLCARQRRLRRRKRKPKATNETDIKKOQLSMPPFKTRKSTDTYSTPAAT 180
DB 121 TQOPTNIRLCARQRRLRRRKRKPKATNETDIKKOQLSMPPFKTRKSTDTYSTPAAT 180
QY 181 TSOPTATYMOCRASDNFSPISPHDRVSTATFAWLHLVQLVLSLQOQLHVQQRSVL 240
DB 181 TSOPTATYMOCRASDNFSPISPHDRVSTATFAWLHLVQLVLSLQOQLHVQQRSVL 240
QY 241 LFRRIAATAFTSYLGSAQKNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 300
DB 241 LFRRIAATAFTSYLGSAQKNSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSSS 300
QY 301 LSAKVCAGVHEKRLHDLDPYNTLRLPVLNDSPLQSLFGLTMOIIVDEKNQLVLT 360
DB 301 LSAKVCAGVHEKRLHDLDPYNTLRLPVLNDSPLQSLFGLTMOIIVDEKNQLVLT 360
QY 361 NVWLKLEWMDNLRWNTSDYGGVKDLRIPPHRIWKPDVLMYNSADEGFGTYQTNVVVRN 420
DB 361 NVWLKLEWMDNLRWNTSDYGGVKDLRIPPHRIWKPDVLMYNSADEGFGTYQTNVVVRN 420
QY 421 NGSCLYVPPGIFKSTCKIDITWPFDDQRCMKFGSWTYDGFQDLQLODETGGDISYV 480
DB 421 NGSCLYVPPGIFKSTCKIDITWPFDDQRCMKFGSWTYDGFQDLQLODETGGDISYV 480
QY 481 LGEWELLGVGKNEIYVNCPPPYDITFPIIRRTLYYFNNLIIPCVLIASMALLG 540
DB 481 LGEWELLGVGKNEIYVNCPPPYDITFPIIRRTLYYFNNLIIPCVLIASMALLG 540
QY 541 FTLPDPSGKLSLGVTLISLTVFLNVAETMPATSDAVPL----- 581
DB 541 FTLPDPSGKLSLGVTLISLTVFLNVAETMPATSDAVPLGTYFNCIMFVASSVVST 600
QY 582 -----WIRIVFLWLPILRMRGRPLILEFPTPCSDTSERKHQ 623
DB 601 ILIINHYHRNADTHEMSEWIRIVFLWLPILRMRGRPLILEFPTPCSDTSERKHQ 660
QY 624 ILSDELKERSKSLANVLDIDDDFRHNCRPMTPGGTLPHPAFYTVYGGDGGSIGP 683
DB 661 ILSDELKERSKSLANVLDIDDDFRHNCRPMTPGGTLPHPAFYTVYGGDGGSIGP 720
QY 684 IGSTRMPDVAVHTHCISKSSEYELGLIKELIRFTDQLRKDDCNDIANDMKFAAMVYDR 743
DB 721 IGSTRMPDVAVHTHCISKSSEYELGLIKELIRFTDQLRKDDCNDIANDMKFAAMVYDR 780
QY 744 LCLIFTMFAIATIAVLLSAPHIIVS 770
DB 781 LCLIFTMFAIATIAVLLSAPHIIVS 807
RESULT 2
Q9VJT9
ID Q9VJT9 PRELIMINARY; PRT; 570 AA.
AC Q9VJT9;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)

DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
GN CG32975-PA.
DN NACRALPHA-34E OR CG32975.
OS Drosophila melanogaster [fruit fly].
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20196006; PubMed=107311132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA Sutton G.G., Mortman J.R., Richards S., Ashburner M., Henderson S.N.,
RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
RA Abril J.F., Agbayani A., An H.J., Andrews-Pfannkoch C., Baldwin D.,
RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferrieria S., Fleischmann W.,
RA Fessler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA Laeko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Malshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D., Scheefer F., Shen H.,
RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun B.,
RA Swirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.Y., Wassarman D.A., Weinstock G.M., Weissbach J.,
RA Williams S.M., Woodagef, Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhang G., Zhao Q., Zheng L.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of Drosophila melanogaster.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Misra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C.M., Berman B.P., Carlson J.W., Celnik S.E.,
RA Clamp M.E., Drysdale R.A., Emmert D., Frise E., de Grey A.D.N.J.,
RA Harris N.L., Kronmiller B., Marshall B., Millburn G.H., Richter J.,
RA Russo S., Searle S.M.J., Smith E., Shu S., Smutniak F.,
RA Whitfield E.J., Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J.,
RA Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE003642; AAP53374.3; -
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.

```
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR006201; Neur_chan.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_memb.
DR Pfam; PF02931; Neur_chan_LBD; 2.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PS00252; NRIONCHANNEL.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
SQ SEQUENCE 570 AA; 64767 MW; DB2E1E1B185263BF CRC64;

Query Match 64.6%; Score 2613; DB 5; Length 570;
Best Local Similarity 69.0%; Pred. No. 1e-219;
Matches 532; Conservative 17; Mismatches 20; Indels 202; Gaps 9;

QY 1 MKNAQLKTEVDDDELWLAHLAHCSSNFSSSSSTRITSSNQRINQQLTTLPQRLSLSTKH 60
DB 1 MKNAQLKTEVDDDELWLAHLAHCSSNFSSSSSTRITSSNQRINQQLTTLPQRLSLSTKH 60
QY 61 HSNIASEQHNSQQOEPASKDEVDVANHGSRNDQOQTHLQQLDSSNMLSPKTAATAAGDEA 120
DB 61 HSNIASEQHNSQQOEPASKDEVDVANHGSRNDQOQTHLQQLDSSNMLSPKTAATAAGDEA 120
QY 121 TQOQTNIRLCARQRORRRRRKPKATNETDIKKQQQLSMPFFKTRKSTDTYSTPAAT 180
DB 121 TQOQTNIRLCARQRORRRRRKPKATNETDIKKQQQLSMPFFKTRKSTDTYSTPAAT 180
QY 181 TSCTPTATYMCRASDNESFIPISRHDRVSTATFAWLVHLVQLLVLSLQQLHVVQORSVL 240
DB 181 TSCTPTATYMCRASDNESFIPISRHDRVSTATFAWLVHLVQLLVLSLQQLHVVQORSVL 240
QY 241 LFRRIAASITAFISYLGSAFAQLKNSSSSSSS-SNSNNSNSTOILNGLNKHSMIFLLIYL 299
DB 241 LFRRIAASITAFISYLGSAFAQLKNSSSSSSS-SNSNNSNSTOILNGLNKHSMIFLLIYL 299
QY 300 NLSAKVCLAGYHEKRLHLDLPYNTLERPVLNESDPLQSLFGLTLMQIIDVDEKNQLLV 359
DB 301 NLSAKVCLAGYHEKRLHLDLPYNTLERPVLNESDPLQSLFGLTLMQIIDVDEKNQLLV 359
QY 360 QLRKDECDNDIANDWKFAAMVVDRLCLIFTMFAIATIAVLSAPHIIVS 770
DB 520 WLRKDECDNDIANDWKFAAMVVDRLCLIFTMFAIATIAVLSAPHIIVS 570

RESULT 3
Q8T5F5
ID Q8T5F5 PRELIMINARY; PRT; 482 AA.
AC Q8T5F5;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
```



```
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Parasas V., Park S., Patel S., Pfeiffer B., Scheeler F.,
RA Phouanavong S., Pittman G.S., Furi V., Richards S., Schaefer F.,
RA Stapleton M., Strong R., Svirska R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Masra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochuk S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Bergman B., Carlson J.W., Celniker S.E.,
RA Clamp M., Drysdale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Kronmiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of Drosophila melanogaster genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celniker S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL: AE003511; AAF48950.2; -; mac-alpha-18C.
DR FlyBase; FBgn0031014; mac-alpha-18C.
DR GO: GO:0016020; C:membrane; IEA.
DR GO: GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
DR GO: GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO: GO:0006811; P:ion transport; IEA.
DR InterPro: IPR006201; Neur_chan.
DR InterPro: IPR006201; Neur_chan_LBD.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR PRINTS: PR00252; NFIONCHANNEL.
DR PROSITE: PS00236; NEUROTR ION CHANNEL; 1.
SQ SEQUENCE 545 AA; 61517 MW; 7B83025107A66209 CRC64;
Query Match 49.2%; Score 1988; DB 5; Length 545;
Best Local Similarity 71.0%; Pred. No. 4.9e-165;
Matches 384; Conservative 46; Mismatches 51; Indels 60; Gaps 8;
276 SNNSSSTQILNGLNKHSMIFLLIYNLSAKVCLAGYHEKRLHLLDPYNTLERPVNLS 335
15 SNNSSSTQILNGLNKHSMIFLLIYNLSAKVCLAGYHEKRLHLLDPYNTLERPVNLS 65
336 PLOLSFGLTLQIIVDEKQKLLVNTVWKLWMDNLRWNTSDYGVKDLRIPPHRIWK 395
66 PLOLSFGLTLQIIVDEKQKLLVNTVWKLWMDNLRWNTSDYGVKDLRIPPHRIWK 125
396 PDVLMYNSADEGDTGYQTNNVVRNNGSLYVPGIFKSTCKIDITWPFDDQRCMKFG 455
126 PDVLMYNSADEGDTGYQTNNVVRNNGSLYVPGIFKSTCKIDITWPFDDQRCMKFG 185
456 SWTYDGFQDLQLODEFGGDISSVVLNGEWELLGVPGKRNIEYVNCCEPYDITFAIII 515
186 SWTYDGFQDLQLODEFGGDISSVVLNGEWELLGVPGKRNIEYVNCCEPYDITFAIII 245
516 RRTLYYFENLIIPCVLIASWALLGFTLPDSGEKISLGVITLLSLTVFLNMVAETMPAT 575
246 RRTLYYFENLIIPCVLIASWALLGFTLPDSGEKISLGVITLLSLTVFLNMVAETMPAT 305
576 SDAVPL-----WIRTVFLCWLPIILRMS 598
306 SDAVPLLGTYNCIMFVASSVSTIILLNHNRPDTHEMSEWIRVIFYLWPCILRMQ 365
599 RGRPLILEFTPTPCSDTSS---ERKQILSDVELKERSKSLIANVLIDDDFRHNCRP 655
366 RPGQ-VGYECPPPPSSSSSSASGEKKQI-QNVELKERSKSLIANVLIDDDFR--CNH 421
```

```
QY 656 MTPGGTLPHPNAPFVRTVYGGDDSGIGPIGST-----RMPDAVTHHTCKSTYEYELG 709
Db 422 RCASATLPHQPTYYRTMYRQDGDGSGVPGPAGPVDGRLHEALS-HTCLTSSAEYELAL 480
QY 710 ILKEIRITDQLRKDDSCNDIANDWKEAAMVDRLCLIIFTMFALLATIAVLLSNPHIIV 769
Db 481 ILKELRWITQLKKEDESDITRDWKEAAMVDRLCLIIFTTIIATLAVLFSAPHFIV 540
QY 770 S 770
Db 541 S 541
RESULT 6
Q9XZ13 PRELIMINARY; PRT; 496 AA.
ID Q9XZ13
AC Q9XZ13
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative nicotinic acetylcholine receptor alpha 7-1 subunit.
OS Heliothis virescens (Noctuid moth) (Owlet moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Heliothinae; Heliothis.
OX NCBI_TaxID=7102;
RN [1]
RP SEQUENCE FROM N.A.
RA Schulte T., Oellers N., Adamczewski M.;
RT "Putative alpha subunits of insect nicotinic acetylcholine receptors
more similar to vertebrate alpha 7 subunits and C. elegans Ce21 than
to other insect nicotinic acetylcholine receptor alpha subunits.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -1- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -1- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL: AF143846; AAD32897.1; -.
DR GO: GO:0016021; C:integral to membrane; IEA.
DR GO: GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
DR GO: GO:0005216; F:ion channel activity; IEA.
DR GO: GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO: GO:0006811; P:ion transport; IEA.
DR GO: GO:0007268; P:synaptic transmission; IEA.
DR InterPro: IPR006201; Neur_chan.
DR InterPro: IPR006202; Neur_chan_LBD.
DR InterPro: IPR006029; Neu Channel memb.
DR Pfam: PF02931; Neur_chan_LBD; 1.
DR Pfam: PF02932; Neur_chan_memb; 1.
DR PRINTS: PR00252; NFIONCHANNEL.
DR TIGRFS: TIGR00860; LIC; 1.
DR PROSITE: PS00236; NEUROTR ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 496 AA; 56347 MW; 8032FED8515A6210 CRC64;
Query Match 45.8%; Score 1850; DB 5; Length 496;
Best Local Similarity 70.0%; Pred. No. 5.1e-153;
Matches 361; Conservative 30; Mismatches 49; Indels 76; Gaps 7;
295 LLIVLNLISAKVCLAGYHEKRLHLLDPYNTLERPVNLSDDPLQSLGTLMOIIVDEX 354
17 LLJLLCLLWFGARGCGHEKRLHLLDHVNLPRPVNNSDDPLQSLGTLMOIIVDEX 76
355 NQLLVNTVWKLWMDNLRWNTSDYGVKDLRIPPHRIWKPDVLMYNSADEGDTGYOT 414
77 NQLLVNTVWKLWMDNLRWNTSDYGVKDLRIPPHRIWKPDVLMYNSADEGDTGYOT 136
415 NVVVRNNGSLYVPGIFKSTCKIDITWPFDDQRCMKFGSWTYDGFQDLQLODEFGG 474
137 NVVVRNNGSLYVPGIFKSTCKIDITWPFDDQRCMKFGSWTYDGFQDLQLODEFGG 196
475 DISSVVLNGEWELLGVPGKRNIEYVNCCEPYDITFAIIIRRTLYYFENLIIPCVLIA 534
```

```
Db 197 DISSFVTNGEWELIGVPGKRNEIYVNCPEPYIDITEAVVIRKTLAYFFENLVPCVLLA 256
Qy 535 SMALLGFTLPDPSGKSLGVTLLSLTVFLNMVAETMPATSDAVPL----- 581
Db 257 SMALLGFTLPDPSGKSLGVTLLSLTVFLNMVAETMPATSDAVPLLGTYFNCIMFWA 316
Qy 582 -----WIRIVFLCWLFWILRMSRPG---RPLILEPPTPCS 614
Db 317 SSVVSTILLNHYHRHADTHMSDWCIRCVFLWLPWLRLMSRPGSATTPPARVPPP-- 374
Qy 615 DTSSERKHQLSDVELKERSKSLANVLDDDDFRHNCRPMTPGGTLPHNPAFYRTVYG 674
Db 375 -----DLELRERSKSLANVLDDDDFRH-----PQAQOPCCRYR---- 412
Qy 675 QGDGSGTIGSTRMPDAVTHHTCKSTSTEVGLGLKEIRFITDQRLKDCNDIANDW 734
Db 413 GGEENGAG-----LAHSCF--GVDYELSLILKEIRVITDQMKKODDADISRDM 460
Qy 735 KFAAMVVDRLCLIIFTFAILATIAVLISAPHIIVS 770
Db 461 KFAAMVVDRLCLIIFTFAILATIAVLISAPHIIVS 496

RESULT 7
Q8T7S2
ID Q8T7S2 PRELIMINARY; PRT; 494 AA.
AC Q8T7S2;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Nicotinic acetylcholine receptor Dalphaf6 subunit variant type II.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
[1]
RN NCBI_TaxID=7227;
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalphas, Dalphaf6 and Dalphaf7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533(2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321446; AAM13393.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:00030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTRP ION CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 494 AA; 56048 MW; 6BE711810ED7B8B CRC64;

Query Match 40.2%; Score 1625.5; DB 5; Length 494;
Best Local Similarity 61.4%; Pred. No. 2.3e-133;
Matches 320; Conservative 47; Mismatches 75; Indels 79; Gaps 7;
Qy 291 SWIFLLIYLNISAKVCLAGYHEKRLHLHLLDLPYNTLFRPVLSNEDPQLSFLTLQMQLID 350
```

```
Db 10 SLFVLLIFLAIKESC--QGPHEKLLNHLSTYNTLERPVANESLEPKFGLTLQIID 68
Qy 351 VDEKNQLLITNVWLKLEWNDMLRWNTSDYGVKDLRI PPHRIWKPDVLYNSADEGPDG 410
Db 69 VDEKNQLLITNLWLSLEWNDYLNKRWNETEYGVKDLRI TPNKWKPDVLYNSADEGPDG 128
Qy 411 TYQTNVVRNNGSLYVPPGIFKSTCKIDITWFFDDQRCQEMKFGSWTYDGFOLDLQIQD 470
Db 129 TYHTNIVVHGSGCLYVPPGIFKSTCKMDITWFFDDQHCQEMKFGSWTYDGNQLDLVLS 188
Qy 471 ETGDDISYVVLNGEWELLGVPGKRNEIYVNCPEPYIDITEAVVIRKTLAYFFENLVPC 530
Db 189 EDGDLSDFITNGEWWYLLAMPGKKNITVYACPEPYVDITFTIQIRRTLYFFENLVPC 248
Qy 531 VLTIASMLLGLTLPDPSGKSLGVTLLSLTVFLNMVAETMPATSDAVPL----- 581
Db 249 VLTISSMLLGLTLPDPSGKSLGVTLLSLTVFLNMVAESMPTSDAVPLIGTYFNCIM 308
Qy 582 -----WIRIVFLCWLFWILRMSRPGSATTPPARVPPP-- 359
Db 309 FVASSVVLTVVVLNHYHRTADIHMPPIKSVFLQWLFWILRMRGPRKI----- 398
Qy 614 SDTSSEKHKQLS-----DVELKERSKSLANVLDDDDFRHNCRPMTPGGTLPHNPAFY 669
Db 360 -----TRKTIILSNRMEKELKERSKSLANVLDDDDFRHTI----- 398
Qy 670 RTVYGQGDGSGTIGSTRMPDAV--THHTCKSTSTEVGLGLKEIRFITDQRLKDCND 728
Db 399 -----SGSQTALGSSASGRTTVEEHTAI--GCNKKDLHLILKELOFITARMKADDEA 452
Qy 729 DIANDWKFAAMVVDRLCLIIFTFAILATIAVLISAPHIIV 769
Db 453 ELIGDWKFAAMVVDRLCLIIFTFAILATIAVLISAPHIIV 493

RESULT 8
Q8T7S3
ID Q8T7S3 PRELIMINARY; PRT; 494 AA.
AC Q8T7S3;
DT 01-JUN-2002 (Tremblrel. 21, Created)
DT 01-JUN-2002 (Tremblrel. 21, Last annotation update)
DE Nicotinic acetylcholine receptor Dalphaf6 subunit variant type I.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
[1]
RN NCBI_TaxID=7227;
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalphas, Dalphaf6 and Dalphaf7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533(2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321445; AAM13392.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:00030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
Qy 291 SWIFLLIYLNISAKVCLAGYHEKRLHLHLLDLPYNTLFRPVLSNEDPQLSFLTLQMQLID 350
```


RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalphas5, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing.";
RT Genetics 160:1519-1533(2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321447; nacr-alpha-30D.
DR FlyBase; FBgn0032151; nacr-alpha-30D.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0005216; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0007268; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 494 AA; 56113 MW; 48327537229573FF CRC64;

Query Match 40.1%; Score 1620.5; DB 5; Length 494;
Best Local Similarity 61.0%; Pred. No. 6.2e-133; Indels 79; Gaps 7;
Matches 316; Conservative 49; Mismatches 75;

QY 291 SWIFLLIYLNLAKVCLAGVHEKRLHLLDPYNTLERPVNLSDPLQLSGLTLMQIID 350
DB 10 SLFVLLIFLAIKESC-QGPEKRLNHLSTYNTLERPVANSEPLEVFGTLQIID 68

QY 351 VDEKNQLLVNTVWMLKLEWDMNLWNTSDYGGVKDLRIPIPHRIWKPDVLMYNSADEGFDG 410
DB 69 VDEKNQLLITNMLSLWMDNLYRNWNETYGGVKDLRIITENKMLKPDVLMYNSADEGFDG 128

QY 411 TYOTNVVRNNGSLYVPGIFKSTCKIDITWFPDDORCEMKFGSWTYDGFQDLQLOD 470
DB 129 TYHTNVVVKSGSLYVPGIFKSTCKMDITWFPDDQHCMEKFGSWTYDGNQDLVNS 188

QY 471 ETGDDISSYVLNGEWELLGVPGKREIYVNCCEPYDITFAIIRRTLYFFNLIPIC 530
DB 189 EDGDLSDFITNGEYLLAMPKKNITVYACCEPYVDITFTQIRRTLYFFNLIVPC 248

QY 531 VLASMAALLGFTLPDPSGEKLSGLVTLLSLTVFLNVAETMPATSDAVPL----- 581
DB 249 VLISSMALLGFTLPDPSGEKLTGLVTLLSLTVFLNLVAESMPTSDAVPLIGVTLLSL 308

QY 582 -----WIRIVFLCWLPLRLMRSPGRPLILEFFPTTC 613
DB 309 FMVASSVVLTVVNLVYHRTADIHEMPWIKSVFLQWLPLRLMRSPGRKI----- 359

QY 614 SDTSSEKXHQILS-----DVELKERSKSLANVLIDDDFRHNCRMPGOTLPHNPAFY 669
DB 360 -----TRKTILLSNRMKELKERSKSLANVLIDDDFRHTI----- 398

QY 670 RTVYGQDDSGISGIPGSTRMPDAV-THHTCTCKSTEYELGLIKELRIFIDQLRKDECN 728
DB 399 -----SGSQAISSAGSSAGFPPTVEEHTAI-GCNHKDLHLILKELOFITARMRKADDEA 452

QY 729 DIANDKFAAMVVDRLCLIIFTFAIATIAVLISAPHIIV 769
DB 453 ELIGDKWFAAMVVDRLCLIVFTLTIATITVLLSAPHIIV 493

RESULT 11

Q8T7R9

ID Q8T7R9

AC Q8T7R9

DT 01-JUN-2002 (TrEMBLrel. 21, Created)

DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalphas6 subunit variant type V.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
[1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes, Dalphas5, Dalphas6 and Dalphas7, in Drosophila melanogaster Identify a New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-Mediated A-to-I Pre-mRNA Editing.";
RT Genetics 160:1519-1533(2002).
RL -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321449; AAM13396.1; -.
DR FlyBase; FBgn0032151; nacr-alpha-30D.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0005216; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan_LBD.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_LBD.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR_ION_CHANNEL; 1.
KW Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
KW Transmembrane.
SQ SEQUENCE 523 AA; 59110 MW; 1C200AF74F87F841 CRC64;

Query Match 39.8%; Score 1611; DB 5; Length 523;
Best Local Similarity 58.2%; Pred. No. 4.6e-132;
Matches 320; Conservative 46; Mismatches 76; Indels 108; Gaps 7;

QY 291 SWIFLLIYLNLAKVCLAGVHEKRLHLLDPYNTLERPVNLSDPLQLSGLTLMQIID 350
DB 10 SLFVLLIFLAIKESC-QGPEKRLNHLSTYNTLERPVANSEPLEVFGTLQIID 68

QY 351 VDEKNQLLVNTVWMLKLEWDMNLWNTSDYGGVKDLRIPIPHRIWKPDVLMYNSADEGFDG 410
DB 69 VDEKNQLITNMLSLWMDNLYRNWNETYGGVKDLRIITENKMLKPDVLMYNSADEGFDG 128

QY 411 TYOTNVVRNNGSLYVPGIFKSTCKIDITWFPDDORCEMKFGSWTYDGFQDLQLOD 470
DB 129 TYHTNVVVKSGSLYVPGIFKSTCKMDITWFPDDQHCMEKFGSWTYDGNQDLVNS 188

QY 471 ETGDDISSYVLNGEWELLGVPGKREIYVNCCEPYDITFAIIRRTLYFFNLIPIC 530
DB 189 EDGDLSDFITNGEYLLAMPKKNITVYACCEPYVDITFTQIRRTLYFFNLIVPC 248

QY 531 VLASMAALLGFTLPDPSGEKLSGLVTLLSLTVFLNVAETMPATSDAVPL----- 581
DB 249 VLISSMALLGFTLPDPSGEKLTGLVTLLSLTVFLNLVAESMPTSDAVPLIGVTLLSL 308

QY 582 -----WIRIVFLCWLPLRLMRSPGRPLILEFFPTTC 613
DB 309 FMVASSVVLTVVNLVYHRTADIHEMPWIKSVFLQWLPLRLMRSPGRKI----- 359

QY 614 SDTSSEKXHQILS-----DVELKERSKSLANVLIDDDFRHNCRMPGOTLPHNPAFY 669
DB 360 -----TRKTILLSNRMKELKERSKSLANVLIDDDFRHTI----- 398

QY 670 RTVYGQDDSGISGIPGSTRMPDAV-THHTCTCKSTEYELGLIKELRIFIDQLRKDECN 728
DB 399 -----SGSQAISSAGSSAGFPPTVEEHTAI-GCNHKDLHLILKELOFITARMRKADDEA 452

QY 729 DIANDKFAAMVVDRLCLIIFTFAIATIAVLISAPHIIV 769
DB 453 ELIGDKWFAAMVVDRLCLIVFTLTIATITVLLSAPHIIV 493

```

Db 415 NVLDIDDDFRHTI-----SGSQTALSSASGRPTTVEEHTAI 453
QY 700 KSSTYELGILKEIRITDOLRKDDNCNDIANDWKFAAMVVDRLCLIIFTMFAILATIA 759
Db 454 -GCNKHDLHLKELQETITARMKADAEALIGDWKFAAMVVDRLCLIVFTLFIATVT 512
QY 760 VLLSAPHIIV 769
Db 513 VLLSAPHIIV 522

RESULT 12
Q9XZ14
ID Q9XZ14 PRELIMINARY; PRT; 501 AA.
AC Q9XZ14;
DT 01-NOV-1999 (TrEMBLrel. 12, Created)
DT 01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Putative nicotinic acetylcholine receptor alpha 7-2 subunit.
OS Heliothis virescens (Noctuid moth) (Owlet moth).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea;
OC Noctuidae; Heliothinae; Heliothis.
OX NCBI_TaxID=7102;
RN [1]
RP SEQUENCE FROM N.A.
RA Schultze T., Oellers N., Adamczewski M.;
RT "Putative alpha subunits of insect nicotinic acetylcholine receptors
RT more similar to vertebrate alpha 7 subunits and C. elegans Ce21 than
RT to other insect nicotinic acetylcholine receptor alpha subunits.";
RL Submitted (APR-1999) to the EMBL/GenBank/DBJ databases.
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF143847; AAD32698.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_memb.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRfams; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
DR Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
DR Transmembrane.
SQ SEQUENCE 501 AA; 56704 MW; 43CB0DC3960C78AB CRC64;

Query Match 39.8%; Score 1609; DB 5; Length 501;
Best Local Similarity 60.8%; Pred. No. 6.4e-132;
Matches 319; Conservative 57; Mismatches 65; Indels 84; Gaps 10;

QY 295 LLYLNLAKVCLAGVHEKRLHDLDPYNTLRLPVLNEDPQLSGLTLMQIIDVDEK 354
Db 10 LLALLPVSEQ---GPHEKRLNALLANYTLRLPVLNEDPQLSGLTLMQIIDVDEK 65
QY 355 NQLLVNTNVLKRWNDNLRWNTSDYGVKDLRTPHRIKWPDLVLYNSADEGFDGTYQT 414
Db 66 NQLLVNTNVLKRWNDNLRWNTSDYGVKDLRTPHRIKWPDLVLYNSADEGFDGTYQT 125
QY 415 NVVVRNNGSLYVPPGIFKSTKIDITWFFDDQRCMKFGSWTYDGFOLDLQDBETGG 474
Db 126 NVVVRSGSLYVPPGIFKSTKIDITWFFDDQRCMKFGSWTYDGNQLDLVLEAGG 185
QY 475 DISYVLNGSWELLGVPGKNEIYNNCCPEFYDITFAIIRRTLYFFENLIIPCVLIA 534
Db 186 DLSDFITNGEWLIGMGKNTITYACCPPEYVDVTFIMIRRTLYFFENLIVPCVLIS 245

```

```

QY 535 SMALLGFTLPDPSGEKLSIGVTLLSLTFLNVAETMPATSDAVPL----- 581
Db 246 SMALLGFTLPDPSGEKLSIGVTLLSLTFLNVAETLPQVSDAIPLLGTGYFNCIMFWVA 305
QY 582 -----WIRIVLCWLPWILRMSRPGRLILBPTTPPCSDTS 617
Db 306 SSVVLTVVLYNHHTADITHMPQWIKSVFLOLWLPWILRMSRPGKKTIRK---TIMNTR 362
QY 618 SERKHQLSDVELKERSKSLANVLDDDDFRHRCRPMTPG-----GTLPHNPAFYRTV 672
Db 363 -----MRELKERSKSLANVLDDDDFRHG--PPPNSTASTGNLPGCSIFRTD 413
QY 673 YQGQ-----DDGSIGPIGSTMPDPAVTHHCISKSTYELGLLIILKEIRFITDQLRKD 724
Db 414 FRKSFVRPSTMEDVG--GGLGS-----ELHLIRELQFIRMKKA 455
QY 725 DECNDIANDWKFAAMVVDRLCLIIFTMFAILATIAVLLSAPHIIV 769
Db 456 DEAEALISDWKFAAMVVDRLCLFVFTLTIATVAVLLSAPHIIV 500

RESULT 13
Q8T7S0
ID Q8T7S0 PRELIMINARY; PRT; 509 AA.
AC Q8T7S0;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Nicotinic acetylcholine receptor Dalpha6 subunit variant type IV.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=21969411; PubMed=11973307;
RA Grauso M., Reenan R.A., Culetto E., Sattelle D.B.;
RT "Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
RT Dalpha5, Dalpha6 and Dalpha7, in Drosophila melanogaster Identify a
RT New and Highly Conserved Target of Adenosine Deaminase Acting on RNA-
RT Mediated A-to-I Pre-mRNA Editing.";
RL Genetics 160:1519-1533(2002).
CC -!- SUBCELLULAR LOCATION: INTEGRAL MEMBRANE PROTEIN (BY SIMILARITY).
CC -!- SIMILARITY: BELONGS TO THE LIGAND-GATED IONIC CHANNEL FAMILY.
DR EMBL; AF321448; AAM13395.1; -.
DR FlyBase; FBgn0032151; nAcr-alpha-30D.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0005216; F:ion channel activity; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR GO; GO:0007268; P:synaptic transmission; IEA.
DR InterPro; IPR006201; Neur_chan.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neur_chan_memb.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRfams; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
DR Glycoprotein; Ionic channel; Postsynaptic membrane; Receptor;
DR Transmembrane.
SQ SEQUENCE 509 AA; 57887 MW; BE8D8E019850C2BD CRC64;

Query Match 39.6%; Score 1602; DB 5; Length 509;
Best Local Similarity 59.3%; Pred. No. 2.7e-131;
Matches 318; Conservative 48; Mismatches 76; Indels 94; Gaps 8;

QY 291 SWIFLLIYLNLAKVCLAGVHEKRLHDLDPYNTLRLPVLNEDPQLSGLTLMQIID 350
Db 10 SLFVLLIYLNLAKVCLAGVHEKRLHDLDPYNTLRLPVLNEDPQLSGLTLMQIID 68

```

QY 351 V-----DEKNOLLVTNVWLKLEWMDMLRWNTSDYGGVKOLRIPPHRIWK 395
DB 69 VDEKNQILTTNAWNLDKKNQLITNLWLSLEWMDYNLRWNETEYGGVKOLRITPNKLNK 128
QY 396 PDVLMYNSADEGFGDTQYVTVVVRNNGSCLYVPGIPKSTCKIDITWPPDPDQCEMKFG 455
DB 129 PDVLMYNSADEGFGDTYHTNVVWRSGSCLYVPGIPKSTCKMIDITWPPDPDQCEMKFG 188
QY 456 SWTYDGFOLDLODETGDISVYVINGEWELLGVGKREIYVNCCEPEYIDITFALII 515
DB 189 SWTYDGNQDLVLNSEDGGLSDFITNGEWMYLLAMPKXNTIVVACPEPYVDTFTFIQI 248
QY 516 RRRLYYFFNLIIPCVLIASMAILGFTLPDPSGKLSIGVTLLSLVFLNVAETMPAT 575
DB 249 RRRLYYFFNLIIPCVLISSWALLGFTLPDPSGKLTGLVTLISLVFLNVAETLPQV 308
QY 576 SDAVPL-----WIRVFLCWLFWILRMS 598
DB 309 SDAIPLLGTGYNCFIMFVASSVLTVVVLNHHRTADITHMPPIKSVFLQWLFWILRMG 368
QY 599 RPRGPLELFEPTTCSSTSERKHQILS---DVELKERSKSLANVLDDDDFRHNCR 654
DB 369 RPRGKI-----TRKILLNSRMKELELERSKSLANVLDDDDFRHTI- 413
QY 655 PMTPGGTLPHPNAPFYRTVYGGDGSIGPIGSTRMPDAV-THHTCIKSSTEYELGLILKE 713
DB 414 -----SGSQTALGSSASFCRPTTVEEHHTAI-GCNHKLHLILKE 452
QY 714 IRFTDOLRKDCENDIANWKFAAMVVDRCLLIFTWFAILATIAVLLSAPHIIV 769
DB 453 LQFITARMRKADDEAELIGDWKFAAMVVDRFCLIVFTLFTIATVTVLLSAPHIIV 508

RESULT 14

Q9NKKD2 ID Q9NKKD2 PRELIMINARY; PRT; 391 AA.
AC Q9NKKD2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN NACR-ALPHA-34E OR BG:DS058999.4 OR BG:DS058999.5 OR CG4498 OR CG16878
GN OR CG32975.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=9403001; PubMed=10471707;
RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA Doyle C., Galle R., George R., Harris N., Hartzell G., Harvey D.,
RA Hong L., Houston K., Hoskins R., Johnson G., Martin C., Moshrefi A.,
RA Palazzolo M., Reese M.G., Spradling A., Tsang G., Wan K., Whitelaw K.,
RA Celniker S., Rubin G.M.;
RA "An exploration of the sequence of a 2.9-Mb region of the genome of
RT Drosophila melanogaster: the Adh region.";
RL Genetics 153:179-219 (1999).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=9403001; PubMed=10471707;
RA Ashburner M., Misra S., Roote J., Lewis S.E., Blazej R., Davis T.,
RA Doyle C., Galle R., George R., Harris N.L., Hoskins R.A.,
RA Houston K.A., Hummasti S.R., Karra K., Kearney L., Kim E., Lee B.,
RA Lewis S., Li P., Lonotan M.A., Mazda P., Moshrefi A.R., Moshrefi M.,
RA Nixon K., Pacleb J.M., Park S., Pfeiffer B., Poon L., Sequeira A.,
RA Sethi H., Snir E., Svirskaas R.R., Wan K.H., Weinburg T., Zhang R.,
RA Zieran L.L., Rubin G.M.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

DR EMBL; AE003408; AAF44854.1; -.
DR FlyBase; FBgn028875; nACR-alpha-34E.
KW Hypothetical protein.
SQ SEQUENCE 391 AA; 43974 MW; E2AB465CF275E8C0 CRC64;
Query Match 37.2%; Score 1502; DB 5; Length 391;
Best Local Similarity 94.4%; Pred. No. 1e-122;
Matches 301; Conservative 5; Mismatches 11; Indels 2; Gaps 2;
QY 1 MKNAQLKLTVDDELWLAVLRLAHCHSNFSSSSSTRSSNQRHNOQLTTLOPRLSTKH 60
DB 1 MKNAQLKLTVDDELWLAVLRLAHCHSNISSSSSSTRSSNKRHNOQLTTLOPRLSTKH 60
QY 61 HSNIASQHNSSQOEPAKDEVDVANHGSRNDQOHLQOLDSSNNMSPKTAATAAGDEA 120
DB 61 HSNIASQHNSSQOEPAKDEVDVANHGSRNDQOHLQOLDSSNNMSPKTAATAAGDEA 120
QY 121 TTOOPTNIRLCARRKQRLRRRRKRPATNETDIKKQQQLSMPFPKTRKSTDTYSTPAAT 180
DB 121 TTOOPTNIRLCARRKQRLRRRRKRPATNETDIKKQQQLSMPFPKTRKSTDTYSTPAAT 180
QY 181 TSCPTATYMOCRASDNFSPISRHRDVRVSTATFAWLHLVLQVLVLSQQQLHVVQORSVL 240
DB 181 TSCPTATYMOCRASDNFSPISRHRDVRVSTATFAWLHLVLQVLVLSQQQLHVVQORSVL 240
QY 241 LFRRIAASTIAFISYLGSAQAQKNSSSSSS-SSNNSNNSSTQILNGINKHSWIFLLIYL 299
DB 241 LFRRIAASTIAFISYLGSAQAQKNSSSSSSSSNNSNNSSTQILNGINKHSWIFLLIYL 300
QY 300 NLSAK-VCLAGYHEKRLH 317
DB 301 NLSAKGECHLNYLECCMQH 319

RESULT 15

Q9VLT79 ID Q9VLT79 PRELIMINARY; PRT; 554 AA.
AC Q9VLT79;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE CG4128 protein.
GN NACR-ALPHA-30D OR NACRALPHA-30D OR BCDNA:GH01410 OR CG4128.
OS Drosophila melanogaster (Fruit fly).
OC Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC Ephydroidea; Drosophilidae; Drosophila.
OX NCBI_TaxID=7227;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Berkley;
RX MEDLINE=20196006; PubMed=10731132;
RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA Ananides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA Abril J.F., Aghayani A., An H.-J., Andrews-Pfannkuch C., Baldwin D.,
RA Balow R.F., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA Beeson K.M., Benos P.V., Berman B.P., Shandari D., Bolshakov S.,
RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA Durbin K.J., Evangelista C.C., Ferraz C., Ferreira S., Fleischmann W.,
RA Folsler C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwan C.,
RA Jallali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,

RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA Shue B.C., Sidenkiamos I., Simpson M., Skupski M.P., Smith T.,
RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA Wang Z.-Y., Wassarman D.A., Weinstein G.M., Weissenbach J.,
RA Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT "The genome sequence of *Drosophila melanogaster*.";
RL Science 287:2185-2195(2000).
RN [2]
RP SEQUENCE FROM N.A.
RA Celnik S.E., Adams M.D., Krommiller B., Wan K.H., Holt R.A.,
RA Evans C.A., Gocayne J.D., Ananides P.G., Brandon R.C., Rogers Y.,
RA Banton J., An H., Baldwin D., Banzon J., Beeson K.Y., Busam D.A.,
RA Carlson J.W., Center A., Champe M., Davenport L.B., Dietz S.M.,
RA Dodson K., Dorsett V., Doup L.E., Doyle C., Dresnek D., Farfan D.,
RA Ferriera S., Frise E., Galle R.F., Garq N.S., George R.A.,
RA Gonzalez M., Houck J., Hoskins R.A., Hostin D., Howland T.J.,
RA Ibegwam C., Jalali M., Kruse D., Li P., Mattei B., Moshrefi A.,
RA McIntosh T.C., Moy M., Murphy B., Nelson C., Nelson K.A., Nunoo J.,
RA Pacleb J., Paragas V., Park S., Patel S., Pfeiffer B.,
RA Phouanavong S., Pittman G.S., Puri V., Richards S., Scheeler F.,
RA Stapleton M., Strong R., Svirskas R., Tector C., Tyler D.,
RA Williams S.M., Zaveri J.S., Smith H.O., Venter J.C., Rubin G.M.;
RT "Sequencing of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Mistra S., Crosby M.A., Matthews B.B., Bayraktaroglu L., Campbell K.,
RA Hradecky P., Huang Y., Kaminker J.S., Prochnik S.E., Smith C.D.,
RA Tupy J.L., Bergman C., Bertman B., Carlson J.W., Celnik S.E.,
RA Clamp M., Derydale R., Emmert D., Frise E., de Grey A., Harris N.,
RA Krommiller B., Marshall B., Millburn G., Richter J., Russo S.,
RA Searle S.M.J., Smith E., Shu S., Smutniak F., Whitfield E.,
RA Ashburner M., Gelbart W.M., Rubin G.M., Mungall C.J., Lewis S.E.;
RT "Annotation of *Drosophila melanogaster* genome.";
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Adams M.D., Celnik S.E., Gibbs R.A., Rubin G.M., Venter C.J.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RA FlyBase;
RL Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AR003626; AAF52817.2; -;
DR FlyBase; FBgn0032151; nACR-alpha-30D.
DR GO; GO:0016020; C:membrane; IEA.
DR GO; GO:0005230; F:extracellular ligand-gated ion channel acti. . .; IEA.
DR GO; GO:0030594; F:neurotransmitter receptor activity; IEA.
DR GO; GO:0006811; P:ion transport; IEA.
DR InterPro; IPR006201; Neur_chan.
DR InterPro; IPR006202; Neur_chan_LBD.
DR InterPro; IPR006029; Neu_chan_LBD; 1.
DR Pfam; PF02931; Neur_chan_LBD; 1.
DR Pfam; PF02932; Neur_chan_memb; 1.
DR PRINTS; PR00252; NRIONCHANNEL.
DR TIGRFAMs; TIGR00860; LIC; 1.
DR PROSITE; PS00236; NEUROTR ION CHANNEL; 1.
SQ SEQUENCE 554 AA; 63236 MW; D6C63EB7F5B32A35 CRC64;

Query Match 32.4%; Score 1310.5; DB 5; Length 554;
Best Local Similarity 47.9%; Pred. No. 1e-105;
Matches 286; Conservative 48; Mismatches 92; Indels 171; Gaps 15;

Search completed: May 7, 2004, 11:38:36

Job time : 52.3701 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - nucleic search, using frame_plus_p2n model

Run on: May 8, 2004, 11:50:34 ; Search time 5376.49 Seconds
(without alignments)
4276.744 Million cell updates/sec

Title: US-09-303-232-2
Perfect score: 4043
Sequence: 1 MKNAQLKITEVDDELWLAV.....MFALLATIAVLVSAPHIIVS 770

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 27513289 seqs, 14931090276 residues
Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+p2n.model -DEV=xlh
-O=/cgm2_1/USPTO_spool/US09303232/runat_07052004_101110_23893/app_query.fasta_1.2261
-DB=EST -QPM=fastcap -SUFFIX=rst -MINMATCH=0.1 -LOOP=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=ptc -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09303232 @CGN 1.1 4577 @runat_07052004_101110_23893 -NCPU=6 -ICPU=3
-NO MMAP -LARGEQUERY -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: em_gss_hum:*
18: em_gss_inv:*
19: em_gss_pln:*
20: em_gss_vrt:*
21: em_gss_fun:*
22: em_gss_mam:*
23: em_gss_mus:*
24: em_gss_pro:*
25: em_gss_rod:*
26: em_gss_pbg:*
27: em_gss_vrl:*
28: gb_gss1:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
C 1	1231	30.4	885	12	BG632919	BG632919 GH16126.3
2	912	22.6	1201	9	AL530299	AL530299 AL530299
3	873.5	21.6	922	13	BU915857	BU915857 AGENCOURT
4	848.5	21.0	2296	14	CD013901	CD013901 90134548
5	837.5	20.7	1436	29	AY402873	AY402873 Homo sapi
6	826.5	20.4	1864	11	AK053497	AK053497 Mus sapi
7	826.5	20.4	2916	11	AK051730	AK051730 Mus muscu
8	826.5	20.4	3126	11	AK080415	AK080415 Mus muscu
9	822.5	20.3	1436	29	AY402875	AY402875 Mus muscu
10	803	19.9	1442	29	AY402876	AY402876 Homo sapi
11	802.5	19.8	1374	29	AY406230	AY406230 Homo sapi
12	798	19.7	1454	29	AY402878	AY402878 Mus muscu
13	797	19.7	1374	29	AY406232	AY406232 Mus muscu
14	797	19.7	4290	11	AK029177	AK029177 Mus muscu
15	791	19.6	1442	29	AY402877	AY402877 Pan trogl
16	787.5	19.5	2010	11	AK080475	AK080475 Mus muscu
17	787	19.5	2940	11	AK034228	AK034228 Mus muscu
18	787	19.5	3230	11	AK083157	AK083157 Mus muscu
19	780	19.3	1781	14	CD013889	CD013889 90138814
20	778.5	19.3	4037	11	AK049722	AK049722 Mus muscu
21	778.5	19.3	4046	11	AK051742	AK051742 Mus muscu
22	770	19.0	908	13	BU149265	BU149265 AGENCOURT
23	757	18.7	4589	11	AK030464	AK030464 Mus muscu
24	748.5	18.5	1374	29	AY406231	AY406231 Pan trogl
25	742.5	18.4	607	9	AI292581	AI292581 GH15518.5
26	729.5	18.0	1436	29	AY402874	AY402874 Pan trogl
27	724.5	17.9	1101	29	CNS0001F	AL058211 Drosophil
28	718.5	17.8	683	12	BM639954	BM639954 170006876
29	718.5	17.8	1007	29	CNS000HU	BM073676 Drosophil
30	711.5	17.6	1466	14	CD013891	CD013891 90139069
31	698.5	17.3	3483	11	AK081254	AK081254 Mus muscu
32	691.5	17.1	1034	13	EX403124	EX403124 BX403124
33	689	17.0	1751	14	CD013905	CD013905 90134608
34	689	17.0	2110	14	CD013902	CD013902 90134640
35	689	17.0	2110	14	CD013903	CD013903 90134855
36	685	16.9	902	29	AY407184	AY407184 Homo sapi
37	683	16.9	2513	11	AK033068	AK033068 Mus muscu
38	682.5	16.9	978	29	CNS0006F9	AL064281 Drosophil
39	677	16.7	902	29	AY407186	AY407186 Mus muscu
40	673.5	16.7	2257	11	AK017571	AK017571 Mus muscu
41	663	16.4	895	13	BM735506	BM735506 BX735506
42	653.5	16.2	615	14	CB149460	CB149460 K-EST0205
43	646.5	16.0	833	14	CB245337	CB245337 UI-M-FYO-
44	643.5	15.9	791	14	CF739825	CF739825 UI-M-HDO-
45	643.5	15.9	2074	11	AK087554	AK087554 Mus muscu

ALIGNMENTS

RESULT 1
BG632919/c
LOCUS
DEFINITION
BG632919 885 bp mRNA linear EST 23-APR-2001
GH16126.3prime GH Drosophila melanogaster head pot2 Drosophila
melanogaster cDNA clone GH16126.3 similar to CG128: FBan0004128
, ion channel, located on: 2L 30D1-30E1;: 04/10/2001, mRNA sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Drosophila melanogaster (fruit fly)
Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;


```
Qy 308 AlaGlyTyrHisGluYsArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGlu 327
Db 187 GCAGCCAGTCCAGAGAGACCTTACAGAGCTGGTCAAGAACTCAATCCCTTGGAG 246
Qy 328 ArgProValLeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGln 347
Db 247 AGGCCCGTGGCCCAATGACTCGCAACCACTACCCCTTACTTCTCCCTGAGCCTCCTGCAG 306
Qy 348 IleIleAspValAspGluYsAsnGlnLeuValThrAsnValTrpLeuYsLeuGlu 367
Db 307 ATCATGGACGTGGATGAGAGAACCAAGTTTAAACCAACCAATTTGGCTGCAATGTCT 366
Qy 368 TrpAsnAspMetAsnLeuArgTrpAsnThrSerAspTyrGlyGlyValYsAspLeuArg 387
Db 367 TGGACAGATCACATTTTACAGTGAATGTGCAGAAATCCAGGGGTGAGACTGTTCGT 426
Qy 388 IleProHisArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGly 407
Db 427 TTCCAGATGCCAGATTTGGAACCAACAGACATTTCTCTATAACAGTGTGTGATGCGC 486
Qy 408 PheAspGlyThrTyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrVal 427
Db 487 TTTGACCCCATTCACATCACTAGCTGTGGTGAATTCCTTGGGCATTCGCCAGTWCCTG 546
Qy 428 ProGlyIlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAsp 447
Db 547 CTTCCAGGCATATTCAGAGCTTCTCTACATCGATGACGTGGTTTCCCTTGAATGTG 606
Qy 448 GlnArgCysGluMetLysPheGlySerThrTyrAspGlyPheGlnLeuAspLeuGln 467
Db 607 CAGCAGCTGCAACAGTGAATTTGGTCTCTTACGAGGCTGGTCTTGGATCTGCAG 666
Qy 468 LeuGlnAspGluThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeu 487
Db 667 ATCCAGAG-----GCAGATATCAGTGGTATATATCCCAATCGAGAATGGACCTA 717
Qy 488 LeuGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIle 507
Db 718 GTGGGAATCCCGCAGAGGAGTGAAGGTTCTATGAGTKCTGCAAGAGCCCTACCCC 777
Qy 508 AspIleThrPheAlaIleIleArgArgArgThrLeuTyrTyrPhePheAsnLeuIle 527
Db 778 GATGTCACTTCACTGAGTGCAGTACGCGCGCAGGAGCGCTCTACTATGGCCTCAACCTGCTG 837
Qy 528 IleProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProAspSer 547
Db 838 ATCCCTGTGTCTATCTCCGCGCCCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 897
Qy 548 -GlyGluYsLeuSerLeu-GlyValThrIleLeuLeuSerLeuThrValPheLeuAsnM 567
Db 898 GGGGAGAGAGATTTCCTKGGGATACAGTYTTACTCTCTTACCGTCTTACGCTGCTGCTGCTG 957
Qy 567 etValIleGluThrMetProAlaThrSerAspAlaValProLeuTrpIleArg----- 584
Db 958 TCGTGGCTGAGATCATCGCGCAACATCGATTCGGTACCAITGAWAGCCAGTCTCTCTCG 1017
Qy 585 -----IleValPheLeuCysTrpLeu 591
Db 1018 CCAGCACCATGATCATCGTGGGCTTTGGGTGGTTG 1052

RESULT 3
BU915857
LOCUS
DEFINITION AGENCOURT 10492745 NICHD_XGC_001 Xenopus laevis cDNA clone
IMAGE:6642638 5', mRNA sequence.
ACCESSION BU915857
VERSION BU915857.1 GI:24097771
KEYWORDS EST.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipoidae; Pipidae;
Xenopodinae; Xenopus.
```

```
REFERENCE 1 (bases 1 to 922)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgap@femail.nih.gov
Tissue Procurement: Martha Rebert, Steven L. Klein, Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM14228 row: f column: 14
High quality sequence stop: 746.
FEATURES
source 1..922
/organism="Xenopus laevis"
/mol_type="mRNA"
/db_xref="taxon:8355"
/clone="IMAGE:6642638"
/tissue_type="oocytes"
/lab_host="DH10B (phage-resistant)"
/clone_lib="NICHD_XGC_001"
/notes="vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI;
Cloned unidirectionally. Primer: Oligo dt. Average insert
size 2.2 kb. Constructed by Life Technologies."
```

Alignment Scores:

Pred. No.:	5,82e-68	Length:	922
Score:	873.50	Matches:	158
Percent Similarity:	75.09%	Conservative:	50
Best Local Similarity:	57.04%	Mismatches:	43
Query Match:	21.61%	Indels:	26
DB:	13	Gaps:	4

US-09-303-232-2 (1-770) x BU915857 (1-922)

```
Qy 352 AspGluYsAsnGlnLeuLeuValThrAsnValTrpLeuYsLeuGluTrpAsnAspMet 371
Db 40 GATGAAAAAATCACTGACTTAAACAACAATATGGTCAATGCTCAATGATGATGATGAT 99
Qy 372 AsnLeuArgTrpAsnThrSerAspTyrGlyValYsAspLeuArgIleProProHis 391
Db 100 TATTTGCAGTGGAAACATGTCTGAATACCCCTGGGGTAAAAAATGTTCTTCAGATGGA 159
Qy 392 ArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThr 411
Db 160 CAGATTGGAAAGCCTGATATCTTCTCTATAACAGTGCAGATGATGATGATGATGATGAT 219
Qy 412 TyrGlnThrAsnValValArgAsnGlySerCysLeuTyrValProProGlyIle 431
Db 220 TTTTACTATAATGACTGTGTGACTCCAGTGCAGTGCAGTGCAGTGCAGTGCAGTGCAGT 279
Qy 432 PheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGlu 451
Db 280 TTTCAAAAGTTCGTGTATACATAGATGCGATGGTTCCTTTTTCAGATATGATGATGATGAT 339
Qy 452 MetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuGlu 471
Db 340 CTGAAGTTGGCTCTGTGGACCTATGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 396
Qy 472 ThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuValPro 491
Db 397 -----TCAGACATATCTCAATATATCGTAATAGAGAGTGGGACCTTGTAGCGTTCTCT 450
Qy 492 GlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPhe 511
Db 451 GGAAGAAAGAGTGAGAAATTACTATGCTCAAGGAAACCGTACCTTGTGATGATGATGATGAT 510
Qy 512 AlaIleIleIleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysVal 531
```


QY 649 PheArgHisAsnCysArg---ProMetThrProGly---GlyThrLeuProHisAsnPro 666
 Db 1436 GACAGATGGGATGTCAGGTCATGTGGCCCTCTGTGGGACCCCTC----- 1483
 QY 667 AlaPheTyrArgThrValTyrGlyGlnGlyAspAspGlySerIleGlyProIleGlySer 686
 Db 1484 -----TGCAGCCAGCCGACCTGACCTCTGGGGCCTCAGGTCCCAAGGCTGAG 1531
 QY 687 ThrArgMetProAspAla-----ValThrHisHisThrCysIleIysSerSerThr 703
 Db 1532 GCTCTGCTGAGGAGGTCAGTCTGCTATCACCCAC----- 1570
 QY 704 GluTyrGluLeuGlyLeuIleLeuLysGluIleArgPheIleThrAspDlnLeuArgLys 723
 Db 1571 -----ATGCAGAGGCACCTGGAAGGTGTCACACTACATTCGCCACACCTGCGGCT 1621
 QY 724 AspAspGluCysAsnAspIleAlaAsnAspTrpLysPheAlaAlaMetValValAspArg 743
 Db 1622 GAGGATGCTGACTCTTCGGTGAAGGAGGACTGGAAGTATGTGCCATGTCATCGACAGG 1681
 QY 744 LeuCysLeuIleIlePheThrMetPheAlaIleLeuAlaThrIleAlaValLeuLeu 762
 Db 1682 ATCTTCCTGCTGCTTTATCATCTGCTGCTTCCTGGGACCATCGGCTCTTCTG 1738

RESULT 5
 LOCUS AY402873 1436 bp DNA linear GSS 15-DEC-2003
 DEFINITION Homo sapiens CHRNA3 gene, VIRTUAL TRANSCRIPT, partial sequence,
 genomic survey sequence.
 ACCESSION AY402873
 VERSION AY402873.1 GI:39759856
 KEYWORDS GSS.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 1436)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Inferring nonneutral evolution from human-chimp-mouse orthologous
 gene trios
 JOURNAL Science 302 (5652), 1960-1963 (2003)
 PubMed 14671302
 REFERENCE 2 (bases 1 to 1436)
 AUTHORS Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarawal,A.,
 Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
 Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
 Adams,M.D. and Cargill,M.
 Direct Submission
 TITLE Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
 Rockville, MD 20850, USA
 COMMENT This sequence was made by sequencing genomic exons and ordering
 them based on alignment.
 FEATURES
 source
 1..1436
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 <1..>1436
 /genes="CHRNA3"
 /locus_tag="HCMI369"
 ORIGIN
 Alignment Scores:
 Pred. No.: 2.04e-64 Length: 1436
 Score: 837.50 Matches: 172
 Percent Similarity: 54.71% Conservative: 95
 Best Local Similarity: 35.25% Mismatches: 158
 Query Match: 20.71% Indels: 63
 Db: 29 Gaps: 5

US-09-303-232-2 (1-770) x AY402873 (1-1436)
 QY 312 GluLysArgLeuLeuHisAspLeuLeuAspP-oTyrAsnThrLeuGluArgProValLeu 331
 Db 21 GAGCACCCTCTATTGAGCGCTGTTGAAGATTACATGATGATCATCGCGCTGTGGCC 80
 QY 332 AsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuVal 351
 Db 81 AACGTGTCTGACCCCATCATCATTCAGAGGTGTCATCTCTCAGCTGCTGAAGGTG 140
 QY 352 AspGluLysAsnGlnLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMet 371
 Db 141 GATGAAGTAAACCCAGATCATGAGACCAACTGTGGCTCAAGCAATCTGGAATGACTAC 200
 QY 372 AsnLeuArgTrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIleProHis 391
 Db 201 AAGCTGAAGTGAACCCCTCTGACTATGTGGGAGAGTTCATCGCTGTCCTGCACAG 260
 QY 392 ArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThr 411
 Db 261 AAGATCTGGAAGCCAGACATGTGCTATAACAATGCTGTGGGGATTTCCAGGTGGAC 320
 QY 412 TyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProGlyIle 431
 Db 321 GACAGACCAAGCGCTTACTCAAGTACACTGGGAGGTGACTTGGATACCTCCGCGCATC 380
 QY 432 PheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGlu 451
 Db 381 TTTAAGAGCTCTCTGTAATCGAGCTGACCTACTTCCGTTTGATTACCAAACTGTACC 440
 QY 452 MetLysPheGlySerThrTrpTyrAspGlyPheGlnLeuAspLeuGlnAspGlu 471
 Db 441 ATGAAGTTCGGTTCCTGCTCTACGATAAGCGGAAATCGATCTGCTGATCGGC--- 497
 QY 472 ThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyValPro 491
 Db 498 TCCTTCCATGAACCTCAAGGACTATTGGGAGAGCGCGGAGTGGGCCATCATCAAGCCCCA 557
 QY 492 GlyLysArgAsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPhe 511
 Db 558 GGCTACAAACACGACATCAAGTACACTGCTGCGAGAGATCTACCCCGCATCACATAC 617
 QY 512 AlaIleIleIleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysVal 531
 Db 618 TCCTGTATCATCGCGCGCTGCGCTTGTCTACACCATCAACCTCATCATCCCTGCTG 677
 QY 532 LeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProAspSerGlyGluLysLeu 551
 Db 678 CT 737
 QY 552 SerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThr 571
 Db 738 ACCCTGTGCATTTCTGTCT 797
 QY 572 MetProAlaThrSerAspAlaValProLeu----- 581
 Db 798 ATCCCTTCCACCTCGCTGCTGCTATCCCGCTGATTGGAGAGTACCTCTCTGTTCCACCATG 857
 QY 581 ----- 581
 Db 858 TTGTGAACCTGTCCATCGTCATCACCGCTTGTGTCCTCAACGTGCATCAGAACCCCG 917
 QY 582 -----TrrIleArgIleValPheLeuCysTrpLeuProTrpIle 594
 Db 918 ACGACACACACAATGCCCTCATGGGTGAAGACTGATTCTTGAACCTGCTCCCGAGGGT 977
 QY 595 LeuArgMetSerArgProGlyArgProLeuIleLeuGluPheProThrThrProCysSer 614
 Db 978 ATG-----TTCATGACCAAGGCCAACAGC 1001
 QY 615 AspThrSerSerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGluArgSer 634
 Db 1002 AACGAGGGCAACGCTCAGAAGCGCGAGGCCCTCTACGGTGGCGGAGCTCTCAAAATCTGAAT 1061

Score: 826.50 Matches: 176
Percent Similarity: 54.04% Conservative: 98
Best local Similarity: 34.71% Mismatches: 166
Query Match: 20.44% Indels: 67
DB: 11 Gaps: 7

US-09-303-232-2 (1-770) x AK053497 (1-1864)

QY 295 LeuLeuIleTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyrHisGluLysArg 314
:::|:::|:::|:::|:::|
Db 127 ATGCTGATGCTGGTGCTGATGCTGTGCAGTGCCAGCGCCTCGAAAGCTGAGCACCG 186

QY 315 LeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProValLeuAsnGluSer 334
:::|:::|:::|:::|:::|
Db 187 CTGTTCCAGTAGCTCTTTGAAGATTACACGAGATCATCCGCCGGTGCTTAACGTGTC 246

QY 335 AspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAspValaspGluLys 354
:::|:::|:::|:::|:::|
Db 247 CATCTGTCAATCCAGTTTGAGGTGTCATGTCTCAGCTGGTGAAGTGGATGAAGTA 306

QY 355 AsnGlnLeuLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAspMetAsnLeuArg 374
:::|:::|:::|:::|:::|
Db 307 AACCAAGATCATGGAACAACACTGTGGCTGAAGCAAACTGGAATGACTACAAGCTGAAA 366

QY 375 TrpAsnThrSerAspTyrGlyValLysAspLeuArgIleProProHisargIleTrp 394
:::|:::|:::|:::|:::|
Db 367 TGGAAACCCTCTGACTACCAAGGGGTGGAGTTTCATCGCAGTCCCTGCAGAGAAGATCTGG 426

QY 395 LysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGlyThrTyrGlnThr 414
:::|:::|:::|:::|:::|
Db 427 AAACGACATCGTGCTTTAACACACCCTGATGGGATTTCCAAGTGGATGACAAAAACC 486

QY 415 AsnValValArgAsnAsnGlySerCysLeuTyrValProProGlyIlePheLysSer 434
:::|:::|:::|:::|:::|
Db 487 AAAGCTCTACTCAAGTACACAGGAGAAGTAGCTTGGATCCCTCCGCCCATCTTTAAGAGC 546

QY 435 ThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCysGluMethLysPhe 454
:::|:::|:::|:::|:::|
Db 547 TCATGCAAAATCGATGTGACCTACTTCCCGTTGACTACCAAAACTGCACCATGAAGTTC 606

QY 455 GlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuGlnAspGluThrGlyGly 474
:::|:::|:::|:::|:::|
Db 607 GGCTCTGTGCTACGACAAGCAAGGCAAGATGACCTGTGCTCCATWGGC---TCCTCCATG 663

QY 475 AspIleSerSertyrValLeuAsnGlyGluTrpGluLeuLeuGlyValProGlyLysArg 494
:::|:::|:::|:::|:::|
Db 664 AACCTCAAGGACTATTGGGAAAGTGGGAGTGGGCCCATCATTAAGCCCCCGGGCTACAAA 723

QY 495 AsnGluIleTyrTyrAsnCysCysProGluProTyrIleAspIleThrPheAlaIleIle 514
:::|:::|:::|:::|:::|
Db 724 CATTGAATAAGTACAACACTGCTGAGGAGATCTACCAAGACATCATCGTACTCGCTATAC 783

QY 515 IleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleIleProCysValLeuIleAla 534
:::|:::|:::|:::|:::|
Db 784 ATTCCGCGCCTCGCGTGTCTACCATCAACCTCATCTTCGTCGCTCTCATCTCC 843

QY 535 SerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGluLysLeuSerLeuGly 554
:::|:::|:::|:::|:::|
Db 844 TTCTCACTGTGCTCGTCTTCTACTCCCTCCGACTGTGGGAGAGGTGACGCTCTGC 903

QY 555 ValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAlaGluThrMetProAla 574
:::|:::|:::|:::|:::|
Db 904 ATCTCGGTGCTCTCCCTCGACGGTCTTTCTCTCGTGTATCCCGAGACCATCCCTTC 963

QY 575 ThrSerAspAlaValProLeu----- 581
:::|:::|:::|:::|
Db 964 ACCTCACTGTCTATCCCCTTGATCGGGAGTACCTCTCTTCACTATGATTTTGTCAAC 1023

QY 581 ----- 581

1024 TTGTCCATCGTCATCACAGCTTTTGTGCTCAACGTGCACTACAGAACTCGACACACAC 1083

QY 582 -----TrpIleArgIleValPheLeuCysTrpLeuProThrIleLeuArgMet 597

Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIVEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861

⁴ The RIKEN Genome Exploration Research Group Phase II Team and the FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)

⁵ The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Group phase I & II team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 2916)

```
QY 575 ThrSerAspAlaValProLeu----- 581
|||||
Db 1040 ACCTCACGTGGTCAATCCCTTGATCGGGAGTACCTCTCTTCACTATGATTTTGTGCACC 1099
581
QY 581
|||||
Db 1100 TTGTCCATCGTCATCACAGTCTTTGTCTCAACGGTCACTACAGAACTCCGACCACACAC 1159
582
QY 582 -----TtpIleArgIleValPheLeuCysTrpLeuProTrrPllLeuArgMet 597
|||||
Db 1160 ACGATGCCCATCTGGGTCAAGCTGTGTTTGAACCTTCTCCCGAGGTCATGTTTATG 1219
598
QY 598 SerArgProGlyArgProLeuLeuGluPheProThrThrProCysSerAspThrSer 617
|||||
Db 1220 ACTAGGCCA-----ACCAAGC 1234
618
QY 618 SerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGluArgSerSerLysSer 637
|||||
Db 1235 ACCGAG-----GAACACGCCCAAGACGAGAACTTCTACGGTGCT 1276
638
QY 638 LeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsnCysArgProMetThr 657
|||||
Db 1277 GAGCTCTCAACCTGAACTGCTTCAGCGCTGCAGACTCCAAAAGCTGCAAGGAAGGCTAC 1336
658
QY 658 Pro-----GlyGlyThrLeuProHisAsnProAlaPheTyrArgThrValTyrGlyGln 675
|||||
Db 1337 CCTGCCAAGATGGACATGTGGCTAT-----TGCCACCCACCGTAGGTTAAAATCTCA 1390
676
QY 676 GlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAspAlaValThrHis 695
|||||
Db 1391 AATTTCAGTGCACACCTCACAAAGAGCTCCAGTCTCTGAGTCTGTGATGCTGTGTGCTCC 1450
696
QY 696 HisThrCysIleLysSerSerThrGluTyrGluLeuGlyLeuIleLeuLysGluIleArg 715
|||||
Db 1451 -----CTGTCTGCTCTCTCACCAGAAATCAAGAAAGCATCAAGAGTGTGAAG 1498
716
QY 716 PheIleThrAspGlnLeuArgLysAspAspGluCysAsnAspIleAlaAsnAspTrpLys 735
|||||
Db 1499 TATATTCTGNAATATGAAGACACAGATGTAGCCAAAGAGATCAAGATGATGGAAG 1558
736
QY 736 PheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrMetPheAlaIleLeu 755
|||||
Db 1559 TATGTTCCATGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1618
756
QY 756 AlaThrIleAlaValLeuLeu 762
|||||
Db 1619 GGAACCTGCAGATTATTTCTG 1639

RESULT 8
AK080415
LOCUS
DEFINITION
Mus musculus 7 days neonate cerebellum cDNA, RIKEN full-length
enriched library, clone:A730007P14 product:NEURONAL NICOTINIC
ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT, full insert sequence.
AK080415
VERSION
AK080415.1 GI:26348538
KEYWORDS
HTC; CAP trapper.
SOURCE
Mus musculus (house mouse)
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1
REFERENCE
Carninci,P. and Hayashizaki,Y.
AUTHORS
TITLE
High-efficiency full-length cDNA cloning
JOURNAL
Meth. Enzymol. 303, 19-44 (1999)
MEDLINE
99279253
PUBMED
10349636
REFERENCE
2
Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K.,
Itoh,M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.
Normalisation and subtraction of cap-trapper-selected cDNAs to
prepare full-length cDNA libraries for rapid discovery of new genes
```

JOURNAL
MEDLINE
PUBMED
REFERENCE
AUTHORS

Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3
Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P.,
Konno,H., Akiyama,J., Nishi,K., Kitsuai,T., Tashiro,H., Itoh,M.,
Sumi,N., Ishii,Y., Nakamura,S., Hazama,M., Nishino,T., Harada,A.,
Yamamoto,R., Matsumoto,H., Sakaguchi,S., Ikegami,T., Kashiwagi,K.,
Fujiwara,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E., Watahiki,M.,
Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsunaga,S., Kawai,J.,
Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A., and Hayashizaki,Y.
RIKEN integrated sequence analysis (RISA) system--384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
The RIKEN Genome Exploration Research Group Phase II Team and the
FANTOM Consortium.
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5
The FANTOM Consortium and the RIKEN Genome Exploration Research
Group Phase I & II Team.
Analysis of the mouse transcriptome based on functional annotation
of 60,770 full-length cDNAs
Nature 420, 563-573 (2002)
6 (bases 1 to 3126)
Adachi,J., Aizawa,K., Akimura,T., Arakawa,T., Bono,H., Carninci,P.,
Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hashizume,W.,
Hayashida,K., Hayatsu,N., Hiramoto,K., Hiraoka,T., Hirozane,T.,
Hori,F., Imotani,K., Ishii,Y., Itoh,M., Kagawa,I., Kasukawa,T.,
Katoh,H., Kawai,J., Kojima,Y., Kondo,S., Konno,H., Kouda,M.,
Koya,S., Kurihara,C., Matsuyama,T., Miyazaki,A., Murata,M.,
Nakamura,M., Nishi,K., Nomura,K., Numazaki,R., Ohno,M., Ohsato,N.,
Okazaki,Y., Saito,R., Saitoh,H., Sakai,C., Sakai,K., Sakazume,N.,
Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T.,
Sogabe,Y., Tagami,M., Tagawa,A., Takahashi,F., Takaku-Akahira,S.,
Takeda,Y., Tanaka,T., Tomaru,A., Toya,T., Yasunishi,A.,
Muramatsu,M. and Hayashizaki,Y.
Direct Submission
Submitted (16-APR-2002) Yoshihide Hayashizaki, The Institute of
Physical and Chemical Research (RIKEN), Laboratory for Genome
Exploration Research Group, RIKEN Genomic Sciences Center (GSC),
RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama,
Kanagawa 230-0045, Japan (E-mail:genome-res@gsc.riken.go.jp,
URL:http://genome.gsc.riken.go.jp/, Tel:81-45-503-9222,
Fax:81-45-503-9216)
cDNA library was prepared and sequenced in Mouse Genome
Encyclopedia Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in RIKEN.
Division of Experimental Animal Research in Riken contributed to
prepare mouse tissues.
Please visit our web site for further details.
URL:http://genome.gsc.riken.go.jp/
URL:http://fantom.gsc.riken.go.jp/.

FEATURES
source

1. .3126
Location/Qualifiers
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="FANTOM DB:A730007P14"
/db_xref="MGI:2407635"
/db_xref="taxon:10090"
/clone="A730007P14"
/tissue_type="cerebellum"
/clone_lib="RIKEN full-length enriched mouse cDNA library"
/dev_stage="7 days neonate"
152. .1666
/note="unnamed protein product; NEURONAL NICOTINIC
ACETYLCHOLINE RECEPTOR ALPHA 3 SUBUNIT (SPTAAL58471,
evidence: FASTY, 99.8%ID, 100%length, match=1497)
putative"

CDS


```

Db      1143 ACT---CCTAGTCTGCAGTCAGCTCCACACGGCAGGCGCTCCCGAGGAGTCCCGGGCTG 1199
Qy      700 LysSerSerThrGluTyr-----GluLeuGlyLeuLeuLeuGlyGluLeuArgPheLe 717
Db      1200 AGGTCTCTGGAGGTTCCGCAAGATCTTACAGGAAGCAATTAGAGGGGCTCAGCTTCATC 1259
Qy      718 ThrAspGlnLeuArgLysAspGluCysAsnAspIleAlaAsnAspTrpLysPheAla 737
Db      1260 GCACAGCATTGGAGAGTGATGATCGAGATCAGATGATGATCGAGAGCTGGAAATTCGTG 1319
Qy      738 AlaMetValValAspArgLeuCysLeuLeuLeuPheThrMetPheAlaIleLeuAlaThr 757
Db      1320 GCAATGGTCGTCGACCGCTGTCTCTGTGGGTTTCGTGATTGTGTGTTCTTCTGGGCACC 1379
Qy      758 IleAlaValLeuLeu 762
Db      1380 ATGGGCTCTCTCTG 1394

RESULT 13
LOCUS   AY406232
DEFINITION Mus musculus CHRNAL gene, VIRTUAL TRANSCRIPT, partial sequence, genomic survey sequence.
ACCESSION AY406232
VERSION   AY406232.1 GI:39762206
KEYWORDS GSS.
SOURCE    Mus musculus (house mouse)
ORGANISM  Mus musculus
          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1374)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
          Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
          Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
          Adams,M.D. and Cargill,M.
          Inferring nonneutral evolution from human-chimp-mouse orthologous
          gene trios
JOURNAL   Science 302 (5652), 1960-1963 (2003)
PUBMED   14671302
REFERENCE 2 (bases 1 to 1374)
AUTHORS   Clark,A.G., Glanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
          Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
          Ferriera,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
          Adams,M.D. and Cargill,M.
          Direct Submission
          Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
          Rockville, MD 20850, USA
          This sequence was made by sequencing genomic exons and ordering
          them based on alignment.
FEATURES
          source
            1..1374
              /organism="Mus musculus"
              /mol_type="genomic DNA"
              /db_xref="taxon:10090"
              <1..>1374
              /gene="CHRNAL"
              /locus_tag="HCM2488"
          gene
          ORIGIN
          Alignment Scores:
          Pred. No.:      8.6e-61      Length:      1374
          Score:         797.00      Matches:     169
          Percent Similarity: 51.97%      Conservative: 95
          Best Local Similarity: 33.27%      Mismatches:  140
          Query Match:      19.71%      Indels:     104
          DB:                29          Gaps:        8

US-09-303-232-2 (1-770) x AY406232 (1-1374)

Qy      291 SerTrpIlePheLeuLeuLeuTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyr 310
Db      10 TCGACTGTCTCTGCTGCTAGGCTC---TGCTCCGCTGGCCTTGTCTGGGCTCCGAA 66

```

```

Qy      311 HisGluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProVal 330
Db      67 CATGAGACGCGTCTGGTGCAAAAGCTCTTTGAAGACTACAGCAGTGTAGTCCGGCCAGTG 126
Qy      331 LeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAsp 350
Db      127 GAGGACCAACCGTGAGATTGTCAAGTCCCGTGGGTCTACAGCTGATCCAGCTTATCAAT 186
Qy      351 ValAspGluLysAsnGlnLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAsp 370
Db      187 GTGGATGAAGTAATACAGATTGTGCAACCAATGTACGTCTGAAACAGCAATGGGTCCAT 246
Qy      371 MetAsnLeuArgTrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIleProPro 390
Db      247 TACAACCTTGAATGCAATCCAGATGACTATGGAGGAGTGAATAAATTCACATCCCTCG 306
Qy      391 HisArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGly 410
Db      307 GAAAGATCTGGCGCGGAGCTGCTCTCTATTAACAACGACGACGGGACTTTGCCATT 366
Qy      411 ThrTyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProProGly 430
Db      367 GTCAAAATTCACCAAGTGTCTCTGGACTACACCGGCCACATCCTGGACACCGCCGCC 426
Qy      431 IlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCys 450
Db      427 ATCTTTAAAGACTACTGTGAGATCATTTGCTACTCATTTCCTTCGTGACGACGACTGTC 486
Qy      451 GluMetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAsp 470
Db      487 AGCATGAAGCTGGCAGCACTGGACCTATGACGGCTCTGTGGTGGCCATAACCCGAAAGT 546
Qy      471 GluThrGlyGlyAspIleSerSerTyrValLeuAsnGlyGluTrpGluLeuGlyVal 490
Db      547 GACCAG---CCCAGCTGAGTAACCTTCATGGAGAGCGGGAGTGGGTGATCAAGAAGCT 603
Qy      491 ProGlyLysArgAsnGluIleTyrTyrAsnCysCysPro---GluProTyrIleAspIle 509
Db      604 CGGGCTGGAAGCACTGGTGTCTTACTCTCTGCTGCCGCCACCTCCCTACTCTGACATC 663
Qy      510 ThrPheAlaIleIleIleArgArgArgThrLeuTyrTyrPhePheAsnLeuIleIlePro 529
Db      664 ACCTACCACATCGCTATGACGCGCTCCCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTT 723
Qy      530 CysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 549
Db      724 TGCTGTCTTCTTCTTCTTCTTAAACAGCTGGTGTCTTCTGCTGCCACGACTCAGGGAG 783
Qy      550 LysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAla 569
Db      784 AAGATGACGCTGAGCATCTCTGTCTTACTGTCTGCTGACCGTGTTCCTTCTGTTGTTGTG 843
Qy      570 GluThrMetProAlaThrSerAspAlaValProLeu----- 581
Db      844 GAGTAATCCCTTCCACCTCCAGCGCTGCCCTGATCGGGAAGTATATGTGTTTTCACC 903
Qy      581 ----- 581
Db      904 ATGGTCTTGTTCATTGGTCCATCATCACCGTTCATCGTCACTCAACACACACCACCGT 963
Qy      582 -----TrpIleArgIleValPheLeuCysTrpLeuPro 592
Db      964 TGCCCGAGCACCCACATCATGCGCGAGTGGTGGGAGGTTTATTCGACACATATCCCA 1023
Qy      593 TrpIleLeuArgMetSerArgProGlyArgProLeuIleLeuGluPheProThrPro 612
Db      1024 AACATCATGTTTTCCTCCCAATGAAGAAGACCA----- 1056
Qy      613 CysSerAspThrSerSerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGlu 632
Db      1057 -----TCCAGAGAT 1065
Qy      633 ArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspPheArgHisAsn 652

```

```

...
1066 AAACAGAGAAAGATTTTACAGAGACATAGATATCTGAC----- 1110
Qy 653 CysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyrArgThrVal 672
Db 1110 ----- 1110
Qy 673 TyrGlyGlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAspAla 692
Db 1111 -----ATCTCTGGGAACCGCGGTCTCT-----CCACCTATG 1140
Qy 693 ValThrHisHisThrCysIleLeuSerSerThrGluTyrGluLeuGlyLeuIleLeuHis 712
Db 1141 GCGTTTCACCTCCGCTGATCAAGCACCT-----GAGTCAAAAGCGCCATCGAG 1191
Qy 713 GluLeuArgPheIleThrAspGlnLeuArgIleAspGluCysAsnAspIleAlaAsn 732
Db 1192 GCGCTGAAGTACATTGCAGAGACCATGAAGTCAGACAGGAGTCCATTAACGCCGCTGAG 1251
Qy 733 AspTrpIlePheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrMetPhe 752
Db 1252 GAAATGGAAGTATGTCCTGATGATGATGATCATCTCCCGGAGTCTTTATGCTGTG 1311
Qy 753 AlaIleLeuAlaThrIleAlaVal 760
Db 1312 TGTCTCATCGGACGCTGCTGTG 1335

AK029177 4290 bp mRNA linear HTC 18-SEP-2003
Mus musculus 0 day neonate head cDNA, RIKEN full-length enriched
library, clone:4831406G09 product:cholinergic receptor, nicotinic,
alpha polypeptide 1 (muscle), full insert sequence.
AK029177
ACCESSION AK029177.1 GI:26325165
VERSION AK029177
KEYWORDS HTC; CAP trapper.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sclurognathi; Muridae; Murinae; Mus.
1 Carninci, P., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
High-efficiency full-length cDNA cloning
Meth. Enzymol. 303, 19-44 (1999)
99279253
10349636
2 Carninci, P. and Hayashizaki, Y.
Normalizing full-length cDNA libraries for rapid discovery of new genes
Genome Res. 10 (10), 1617-1630 (2000)
20499374
11042159
3 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
Konno, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M.,
Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A.,
Yamamoto, R., Matsumoto, S., Sakaguchi, S., Ikegami, I., Kashiwagi, K.,
Fujiwara, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M.,
Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J.,
Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system-384-format
sequencing pipeline with 384 multicapillary sequencer
Genome Res. 10 (11), 1757-1771 (2000)
20530913
11076861
4
5
The RIKEN Genome Exploration Research Group Phase II Team and the
PANTOM Consortium
Functional annotation of a full-length mouse cDNA collection
Nature 409, 685-690 (2001)
5

```

AUTHORS

TITLE

JOURNAL

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

FEATURES

source

CDS

polya_signal

polya_site

ORIGIN

Alignment Scores:

Pred. No.: 4,93e-60 Length: 4290

Score: 797.00 Matches: 169

Percent Similarity: 51.97% Conservative: 95

The FANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team.

Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs

Nature 420, 563-573 (2002)

6 (bases 1 to 4290)

Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Konno, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, F., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohmoto, N., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y.

Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan [E-mail: genome-res@gsc.riken.go.jp, URL: http://genome.gsc.riken.go.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216]

cDNA library was prepared and sequenced in Mouse Genome

Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.go.jp/

URL: http://fantom.gsc.riken.go.jp/

Location/Qualifiers

1..4290

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM,DB:4831406G09"

/db_xref="MGI:2391229"

/db_xref="taxon:10090"

/clone="4831406G09"

/tissue type="head"

/clone lib="RIKEN full-length enriched mouse cDNA library"

/dev stage="0 day neonate"

21..1394

/note="unnamed protein product; cholinergic receptor,

nicotinic, alpha polypeptide 1 (muscle) [MGI:87885,

GB|NM_007389, evidence: BLASTN, 99%, match=1763]

putative"

/codon start=1

/protein_id="BAC26337.1"

/db_xref="GI:26325165"

/translation="MELSTVLLLLGLCSAGLVGSEHETRLVAKLFEDYSSVVRVED
HREIVQTVGLQIQLINVDVNVQIVTNRLKQWVDYNLKNWDDDDGGVKKIHP
EKIMPLDVTYNNADGDAIYVKFTGLDYTGHTTTPAIFPSYCEIIVTHFPDEQ
NCMKLPGVYDGSVAINPESQDPLDNFMESEGWIKFGEAGWVYSCCTTPY
LDITVHFVMDRLPLFVIVNIIICLLFSLFVLFVLTSGEKMFLSISLSTVF
LLVIVLIPSTSSAVPLIKGMFLTMVFVIAIITIVINTHHSPSTHMPWVRK
VFIDIPNMFSTWKRSDRQKREITEDIDISDKGPGPPGPHSPILKHPEV
KELSGVKYIATMTKSDQESNNAEENKIVAMVMDHILLGVFNLVCLIGTLAVFAGRL
1SLHQHG"

4267..4272

/note="putative"

4290

/note="putative"

Best Local Similarity: 33.27% Mismatches: 140
 Query Match: 19.71% Indels: 104
 DB: 11 Gaps: 8

US-09-303-232-2 (1-770) x AK029177 (1-4290)

```

QY 291 SerTrpIlePheLeuLeuIleTyrLeuAsnLeuSerAlaLysValCysLeuAlaGlyTyr 310
DB 30 TCGACTGTTCTCTGCTAGGCTC--TGCTCCGCTGGCTTGTCTGGGCTCCGAA 86
QY 311 HisGluLysArgLeuLeuHisAspLeuLeuAspProTyrAsnThrLeuGluArgProVal 330
DB 87 CATGAGACGCTCTGGTGCAAGCTCTTTGAAGACTACAGACGTGTAGTCGGCCACTG 146
QY 331 LeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeuThrLeuMetGlnIleLeuAsp 350
DB 147 GAGACACACCTGAGATTGTACAGTCACTCCGCTGCTACAGCTGATCCAGCTTATCAAT 206
QY 351 ValAspGluLysAsnGlnLeuValThrAsnValTrpLeuLysLeuGluTrpAsnAsp 370
DB 207 GTGATGAAGTAATCAGATTGTGACACCAATGTACGTCTGAACACCAATGGTGCAT 266
QY 371 MetAsnLeuArgTrpAsnThrSerAspTyrGlyGlyValLysAspLeuArgIleProPro 390
DB 267 TACAACCTGGAATGGAATCCAGATCACTATGGAGGAGTGAATAAATTTCATCCCTCG 326
QY 391 HisArgIleTrpLysProAspValLeuMetTyrAsnSerAlaAspGluGlyPheAspGly 410
DB 327 GAAAGATCTGGCGCGGACGTCGTCTCTATTAACAACGACGACGGCGACTTTGGCCATT 386
QY 411 ThrTyrGlnThrAsnValValArgAsnAsnGlySerCysLeuTyrValProProGly 430
DB 387 GTCAAAATCCACCAAGTGCTCTGAGTACACCGGCCACATCACTCGACACCCGACCC 446
QY 431 IlePheLysSerThrCysLysIleAspIleThrTrpPheProPheAspAspGlnArgCys 450
DB 447 ATCTTTAAAGACTACTGTGATCATTTGTCTACTCCTTCCCTTCGATGACGAACTGC 506
QY 451 GluMetLysPheGlySerTrpThrTyrAspGlyPheGlnLeuAspLeuGlnLeuAsp 470
DB 507 AGCATGAGCTGGCCACTCGACTATGACGGCTCTGTGTGGCCATTAACCCGGAAGT 566
QY 471 GluThrGlyGlyAspIleSerTyrValLeuAsnGlyGluTrpGluLeuLeuGlyVal 490
DB 567 GACCAG---CCCGACTGAGTAATCTCATGAGAGCGGGAGTGGGTGATCAAGGAAGCT 623
QY 491 ProGlyLysArgAsnGluLysTyrTyrAsnCysPro---GluProTyrIleAspIle 509
DB 624 CGGGGCTGGAAGCACTGGGTGTCTACTCTGCTGCCCCACCACTCCCTACCTGGACATC 683
QY 510 ThrPheAlaIleIleLeuArgArgArgThrLeuTyrTyrPhePheAsnLeuIlePro 529
DB 684 ACTACCACTTCGTATGAGCGCTCGCCCTCTACTTATTGTCACAGTCATCATTTCCC 743
QY 530 CysValLeuLeuAlaSerMetAlaLeuLeuGlyPheThrLeuProProAspSerGlyGlu 549
DB 744 TGCTGCTCTCTCTCTTAAACAGCGCTGGTGTCTACTCGCCACAGACTCAGGGGAG 803
QY 550 LysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrValPheLeuAsnMetValAla 569
DB 804 AGATGACGCTGAGCATCTCTGTCTACTGTCTCCTGACCGTGTCTCTGTTGGTCAATTG 863
QY 570 GluThrMetProAlaThrSerAspAlaValProLeu----- 581
DB 864 GAGCTAATCCCTTCCACCTCCAGCGCTGTGCCCTGATCGGGAAGTATATGTTGTTCCAC 923
QY 581 ----- 581
DB 924 ATGGTCTTTGTCATTGGTCCATCATCACCGTCATCGTCAATCAACACACACCAACCGT 983
QY 582 -----TrpIleArgIleValPheLeuCysTrpLeuPro 592
DB 984 TCGCCGAGCACCACATCATGCGCGAGTGGTGGGGAAGTCTTTTATCGACACTATCCCA 1043

```

```

QY 593 TrpIleLeuArgMetSerArgProGlyArgProLeuLeuLeuGluPheProThrThrPro 612
DB 1044 AACATCATGTTTCTCCACAATGAAGAAGACCA-----TCCAGAGAT 1076
QY 613 CysSerAspThrSerSerGluArgLysHisGlnIleLeuSerAspValGluLeuLysGlu 632
DB 1077 -----TCCAGAGAT 1085
QY 633 ArgSerSerLysSerLeuLeuAlaAsnValLeuAspIleAspAspAspPheArgHisAsn 652
DB 1086 AAAACAAGAGAAAAGATTTTACAGACACATAGATATATCTGAC----- 1130
QY 653 CysArgProMetThrProGlyGlyThrLeuProHisAsnProAlaPheTyrArgThrVal 672
DB 1130 ----- 1130
QY 673 TyrGlyGlnGlyAspAspGlySerIleGlyProIleGlySerThrArgMetProAspAla 692
DB 1131 -----ATCTCTGGGAAGCCGGTCTCT-----CCACCTATG 1160
QY 693 ValThrHisHisThrCysIleLysSerSerThrGluTyrGluLeuGlyLeuIleLys 712
DB 1161 GGCTTTCACTCTCGCTGATCAAGCACCT-----GAGGTGAAGAGCGGCTGAG 1211
QY 713 GluIleArgPheIleThrAspGlnLeuArgLysAspAspGluCysAsnAspIleAlaAsn 732
DB 1212 GGGTGAAGTACATTCAGAGACCATGATGAAGTCAGACGAGAGTCCATATACCGCGCTGAG 1271
QY 733 AspTrpLysPheAlaAlaMetValValAspArgLeuCysLeuIleIlePheThrMetPhe 752
DB 1272 GAATGGAAGTATGTGCTGATGATGATCATCTCTCGAGTCTTTATGCTGTG 1331
QY 753 AlaIleLeuAlaThrIleAlaVal 760
DB 1332 TGCTCATCGGAGCGTGGCTGTG 1355

RESULT 15
AY402877 1442 bp DNA linear GSS 15-DEC-2003
LOCUS Pan troglodytes CHRN4 gene, VIRTUAL TRANSCRIPT, partial sequence,
DEFINITION genomic survey sequence.
ACCESSION AY402877
VERSION AY402877.1 GI:39758860
SOURCE GSS.
ORGANISM Pan troglodytes (chimpanzee)
Pan troglodytes
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
REFERENCE 1 (bases 1 to 1442)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Inferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
Science 302 (5652), 1960-1963 (2003)
JOURNAL PUBMED 14671302
REFERENCE 2 (bases 1 to 1442)
Clark,A.G., Gnanowski,S., Nielson,R., Thomas,P., Kejariwal,A.,
Todd,M.A., Tanenbaum,D.M., Civello,D.R., Lu,F., Murphy,B.,
Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Sninsky,J.J.,
Adams,M.D. and Cargill,M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
Location/Qualifiers
1..1442
/organism="Pan troglodytes"
/mol_type="genomic DNA"
/db_xref="taxon:9598"

```

```
Gene      <1. >1442
           /gene="CHRNA4"
           /locus_tag="HCM1370"

ORIGIN

Alignment Scores:
Pred. No.:      3,22e-60      Length:      1442
Score:          791.00        Matches:      177
Percent Similarity: 51.29%    Conservative: 82
Best Local Similarity: 35.05%  Mismatches:   152
Query Match:      19.56%     Indels:       94
DB:              29          Gaps:        13

US-09-303-232-2 (1-770) x AY402877 (1-1442)

Qy      306 CysLeuAlaGlyTyrHisGluLeuLeuHisAspLeuLeuLeuPro-----Tyr 323
Db      6  TGGCGGTGGCCAAATGGCGAGGAAAAAGCTGATGAGCAGACCTTCTGAACAAAACCGTTAC 65

Qy      324 AsnThrLeuGluArgProValLeuAsnGluSerAspProLeuGlnLeuSerPheGlyLeu 343
Db      66 AATAACCTGATCCGCCAGCCAGCCAGCTCCTCAGCTCATCTCTCAATCAAGCTGAGTGC 125

Qy      344 ThrLeuMetGlnIleleAspValAspGluLysAsnGlnLeuLeuValThrAsnValTrp 363
Db      126 TCCCTGGCCCACTATCAGCGTGAATGAGCGAGAGCAGATCATGACCACCAATGTCTGG 185

Qy      364 LeuLysLeuGluTrpAsnAspMetAsnLeuArgTrpAsnThrSerAspTyrGlyVal 383
Db      186 CTGAACAGGAATGGACTGATTACCGCTGACCTGGAAACAGCTCCCGTACGAGGGTGTG 245

Qy      384 LysAspLeuArgIleProHisArgIleTrpLysProAspValLeuMetTyrAsnSer 403
Db      246 AACATCCTGAGATCCCTCGAAAGCGCACTGCTGTCTACACCACTTGTAGTCCGGTCCAAATGGCAGC 305

Qy      404 AlaAspGluGlyPheAspGlyThrTyrGlnThrAsnValValArgAsnAsnGlySer 423
Db      306 GCCGAGCGGACTATGAGGTGCTGTCTACACCACTTGTAGTCCGGTCCAAATGGCAGC 365

Qy      424 CysLeuTrpValProGlyIlePheLysSerThrCysLysIleAspIleThrTrpPhe 443
Db      366 CTCCTGTGGTCCCGCTGCCATCTACAGAGCGCTGCAAGATTGAGGTGAAGTACTTT 425

Qy      444 ProPheAspAspGlnArgCysGluMetLysPheGlySerTrpThrTyrAspGlyPheGln 463
Db      426 CCTTTCGACGAGCAACTGCACCTCAAGTTCCCTCTCGGACCTATGACCACGCGAG 485

Qy      464 LeuAspLeuGlnLeuAspGluThrGlyGlyAspIleSerSerTyrValLeuAsnGly 483
Db      486 ATAGACATGGTCTTCATGAGGCCACACA--GCCAGCATGGATGACTTTACTCCCAAGTGT 542

Qy      484 GluTrpGluLeuLeuGlyValProGlyLysArgAsnGluIleTyrTyrAsnCysPro 503
Db      543 GAGTGGGACATAGTGGCCCTCCCGAGGAGAGACA-----GTGAACCCACAGGAC 593

Qy      504 GluProTyrIleAspIleThrPheAlaIleIleArgArgThrLeuTyrTrpPhe 523
Db      594 CCCAGCTACGAGAGTGAATTCAGCTTCATCATCAGCGCAAGCCCTGTGTCTACACC 653

Qy      524 PheAsnLeuIleProCysValLeuIleAlaSerMetAlaLeuLeuGlyPheThrLeu 543
Db      654 ATCAACCTCATATCCCTCGTGTCTACCACTGTGTGGCCATCCTTGTCTTCTACCTG 713

Qy      544 ProProAspSerGlyGluLysLeuSerLeuGlyValThrIleLeuLeuSerLeuThrVal 563
Db      714 CCATCCAACTGCGCGGAGAGATGACACTGTGCACTCTGAGTGTGCTGCGCATTTACATTC 773

Qy      564 PheLeuAsnMetValAlaGluThrMetProAlaThrSerAspAlaValProLeu----- 581
Db      774 TTCCTGCTCATCTCCANGNTGCGGCCACCCACCTCCCTCGATGTGCTCTCATCGGC 833

Qy      581 ----- 581
```

Search completed: May 9, 2004, 04:19:59
Job time : 5420.49 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 12:59:49 ; Search time 5894.18 Seconds
(without alignments)
10942.062 Million cell updates/sec

Title: US-09-303-232-3_COPY_335_1822
Perfect score: 1488
Sequence: 1 atggcgggcgggcgccgcg.....cgccacacatcatgtgtcgcg 1488

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 3470272 seqs, 21671516995 residues

Total number of hits satisfying chosen parameters: 6940544

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_htg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vi.*

15: em_ba.*

16: em_fun.*

17: em_hum.*

18: em_in.*

19: em_mu.*

20: em_om.*

21: em_or.*

22: em_ov.*

23: em_pat.*

24: em_ph.*

25: em_pl.*

26: em_ro.*

27: em_sts.*

28: em_un.*

29: em_vi.*

30: em_htg_hum.*

31: em_htg_inv.*

32: em_htg_other.*

33: em_htg_mus.*

34: em_htg_pln.*

35: em_htg_rod.*

36: em_htg_mam.*

37: em_htg_vrt.*

38: em_sy.*

39: em_htgo_hum.*

40: em_htgo_mus.*

41: em_htgo_other.*

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	1488	100.0	3629	3	AF143846	AF143846 Heliothis
2	1488	100.0	3700	6	AX009612	AX009612 Sequence
3	1488	100.0	3701	6	E58347	E58347 Nucleic aci
4	615	41.3	2023	3	AF321446	AF321446 Drosophil
5	611.8	41.1	1665	3	DME554209	AF554209 Drosophil
6	611.8	41.1	1699	3	BT011147	BT011147 Drosophil
7	611.8	41.1	2023	3	AF321447	AF321447 Drosophil
8	608.6	40.9	2023	3	AF321445	AF321445 Drosophil
9	583.8	39.2	2907	3	AF272778	AF272778 Drosophil
10	575	38.6	3029	3	AF143847	AF143847 Heliothis
11	575	38.6	3109	6	E58348	E58348 Nucleic aci
12	575	38.6	3109	6	AX009614	AX009614 Sequence
13	568.4	38.2	1683	3	DME554210	AF554210 Drosophil
14	553.6	37.2	2068	3	AF321448	AF321448 Drosophil
15	519.6	34.9	2110	3	AF321449	AF321449 Drosophil
16	512.8	34.5	2886	6	E58346	E58346 Nucleic aci
17	512.8	34.5	2886	6	AX009610	AX009610 Sequence
18	455.4	30.6	2834	3	AY036613	AY036613 Drosophil
19	410.6	27.6	1848	10	MUSNARS	L37663 Mus musculu
20	410	27.6	2107	9	AF486623	AF486623 Macaca mu
21	407.4	27.4	2088	10	AF225980	AF225980 Mus muscu
22	405.6	27.3	1509	6	AX054567	AX054567 Sequence
23	405.6	27.3	1876	6	AR282833	AR282833 Sequence
24	405.6	27.3	1876	6	AX171908	AX171908 Sequence
25	405.6	27.3	1876	9	HSU62436	U62436 Human nicot
26	405.2	27.2	1876	6	AR055255	AR055255 Sequence
27	405.2	27.2	1876	6	AR071403	AR071403 Sequence
28	405.2	27.2	1876	6	AR171187	AR171187 Sequence
29	405.2	27.2	1876	6	AR224030	AR224030 Sequence
30	404	27.2	1509	6	AX054577	AX054577 Sequence
31	404	27.2	1509	6	AX054577	AX054577 Sequence
32	404	27.2	1509	9	HSNACHRA7	Y08420 H.sapiens m
33	404	27.2	3030	10	S53987	S53987 nicotinic r
34	402.4	27.0	1509	6	AX054579	AX054579 Sequence
35	401.8	27.0	2106	10	RATNARAD	U31619 Rattus ratt
36	400.8	26.9	1916	9	AF385585	AF385585 Homo sapi
37	400.8	26.9	1977	9	HSU40583	U40583 Human alpha
38	400.8	26.9	2087	9	HSARA7A	X70297 H.sapiens m
39	397.6	26.7	1555	9	HSCRNA7A	Z23141 H.sapiens C
40	396.4	26.6	1551	4	BT17	X93604 B.taurus mr
41	396	26.6	1559	9	HUMA7NAR	L25827 Human a7 ni
42	394	26.5	1590	6	AR261850	AR261850 Sequence
43	394	26.5	1590	6	BD023656	BD023656 Variant h
44	364	24.5	1668	5	AY298752	AY298752 Takifugu
45	357.6	24.0	1542	5	AY299466	AY299466 Takifugu

ALIGNMENTS

RESULT 1
AF143846
LOCUS AF143846
DEFINITION Heliothis virescens putative nicotinic acetylcholine receptor alpha
7-1 subunit mRNA, complete cds.
ACCESSION AF143846
VERSION AF143846.1
KEYWORDS GI:4895004
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1 (bases 1 to 3629)
AUTHORS Schulte,T., Oellers,N. and Adamczewski,M.

Pred. No. is the number of results predicted by chance to have a

TITLE Putative alpha subunits of insect nicotinic acetylcholine receptors more similar to vertebrate alpha 7 subunits and C. elegans Ce21 than to other insect nicotinic acetylcholine receptor alpha subunits

JOURNAL Unpublished
REFERENCE 2 (bases 1 to 3629)
AUTHORS Schulte, T., Oellers, N. and Adamczewski, M.
TITLE Direct Submission
JOURNAL Submitted (19-APR-1999) ZF-BTB, Bayer AG, Bldg. Q 18., Leverkusen 51368, Germany

FEATURES
Location/Qualifiers
1..3629
/organism="Heliothis virescens"
/mol_type="mRNA"
/db_xref="taxon:7102"
335..1825
/note="hvnachra7-1"
/codon_start=1
/product="putative nicotinic acetylcholine receptor alpha 7-1 subunit"
/protein_id="AAD32697.1"
/db_xref="GI:4895005"
/translation="MGRARRSHLAAPAGLLLLCLLWPRGARQYHEKRLHLLHLDH YNVLPRVNESDPLQSLFGLTLMQI IDVDEKNQLLIINWLEWMDNLRWNTSDF GGVKDLRVPRLMKFQSVTYDNLQLODEGGDISLVPVNGWELIGVPGKRNEL IYVCCPEPYDIDFPAVIRKTLVFNLIJPCVLJASMLLGLFTLPDPSGEKLSLG VTLLSLVTLNMAETPATSADVPLLLGTTFNCIMENVASVSTILLNHYRHAD THDSMDTRCVFLVLPWLRMSRPSATPPPARVPPDLERKRSKLLANVLIDDDFRHPQACQPCRYRGEENAGLAHSCFVNDYELSLKLEIRVITDQWRKD DEDADISRDWKFAMVVDRLCLLIPTLTITIAVLALLSAPHIMVS"

CDS
100.0%; Score 1488; DB 3; Length 3629;
Best Local Similarity 100.0%; Pred. No. 7.2e-291;
Matches 1488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

ORIGIN
Query Match 100.0%; Score 1488; DB 3; Length 3629;
Best Local Similarity 100.0%; Pred. No. 7.2e-291;
Matches 1488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 ATGGCGGGGCGGGCGCGCTCGACATTGGCGGGCGCGGGCGCTGCTGCTGCTG 60
DB 335 ATGGCGGGGCGGGCGCGCTCGACATTGGCGGGCGCGGGCGCTGCTGCTG 394
61 TGCCTGCTCTGCGCGGGGGGCGACGCTGCGGGTACACAGAGCGGCTACTGCAC 120
DB 395 TGCCTGCTCTGCGCGGGGGGCGACGCTGCGGGTACACAGAGCGGCTACTGCAC 454
121 CTATTTGGACCACTAACGCTACTGAGAGGCGGCTGCTCAACGAGAGCGGCTG 180
DB 455 CTATTTGGACCACTAACGCTACTGAGAGGCGGCTGCTCAACGAGAGCGGCTG 514
181 CTCTCTCTGCGGCTCACGCTCATGCGATCATCGACGCTGAGAGAGAACAGCTTT 240
DB 515 CTCTCTCTGCGGCTCACGCTCATGCGATCATCGACGCTGAGAGAGAACAGCTTT 574
241 ATAAACAAACATCTGCTAAACATAGAGTGAATGATGAACTTGAAGTGAACACTT 300
DB 575 ATAAACAAACATCTGCTAAACATAGAGTGAATGATGAACTTGAAGTGAACACTT 634
301 GATTTTCGGGGGTCAAGATTTAAGGTGTCACCCACAGACTATGGAACACGACGTC 360
DB 635 GATTTTCGGGGGTCAAGATTTAAGGTGTCACCCACAGACTATGGAACACGACGTC 694
361 CTATTTGTAACACAGCGCGGACGAAGGTTTCGACAGCAGCAGTATCCAAACGCT 420
DB 695 CTATTTGTAACACAGCGCGGACGAAGGTTTCGACAGCAGCAGTATCCAAACGCT 754
421 CGGAACAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 480
DB 755 CGGAACAAACGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 814
481 GACATCACTGGTTCCTCTCGACGACCAACGATGCGAGATGAAGTTTGGAGCTG 540
DB 815 GACATCACTGGTTCCTCTCGACGACCAACGATGCGAGATGAAGTTTGGAGCTG 874

RESULT 2
AX009612
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
AX009612
Sequence 3 from Patent EP0962528.
AX009612
AX009612.1 GI:9996844
Heliothis virescens (tobacco budworm)

541 TATGATGTTTATCAGTTGGATCTACAACCTACAGGATGAAGGGGGGAGATATAGCAGT 600
DB 875 TATGATGTTTATCAGTTGGATCTACAACCTACAGGATGAAGGGGGGAGATATAGCAGT 934
601 TTTGTACGAATGGCGAATGGAGTTAATAGGAGTCCCGGCAAGCGCAAGAGATCTAC 660
DB 935 TTTGTACGAATGGCGAATGGAGTTAATAGGAGTCCCGGCAAGCGCAAGAGATCTAC 994
661 TACAACCTGTTTCTCGGAGCCATACATCGACATCACGTTTTCGGTGGTGCATCCGAGGAAA 720
DB 995 TACAACCTGTTTCTCGGAGCCATACATCGACATCACGTTTTCGGTGGTGCATCCGAGGAAA 1054
721 ACGCTCTACTACTTCTTCAATCTCATCGTGCCTTCGCTGCTCATCGCTTCATGCTCTTA 780
DB 1055 ACGCTCTACTACTTCTTCAATCTCATCGTGCCTTCGCTGCTCATCGCTTCATGCTCTTA 1114
781 TTGGGGTTTACCTTTCCTCAGATCTCGGAGAAAAGTTGTTCTTTAGGTGTGACGATATTA 840
DB 1115 TTGGGGTTTACCTTTCCTCAGATCTCGGAGAAAAGTTGTTCTTTAGGTGTGACGATATTA 1174
841 CTGTCTGTTGACGCTGTTTCTTCAACATGTTGGCGGAGACGATGCGAGCAGCTCGGACGCC 900
DB 1175 CTGTCTGTTGACGCTGTTTCTTCAACATGTTGGCGGAGACGATGCGAGCAGCTCGGACGCC 1234
901 GTGCCCTTGTTCGGACACTACTTCAACTGATCATGTTTCATGTTGGTCTTCTTCCGTGTC 960
DB 1235 GTGCCCTTGTTCGGACACTACTTCAACTGATCATGTTTCATGTTGGTCTTCTTCCGTGTC 1294
961 TCCACCATACTGATCTTCAACTACACACCGGAGCAGACACTCAGGAAATGATGAT 1020
DB 1295 TCCACCATACTGATCTTCAACTACACACCGGAGCAGACACTCAGGAAATGATGAT 1354
1021 TGGATTTCGTTGCTTCTTATTTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1080
DB 1355 TGGATTTCGTTGCTTCTTATTTGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1414
1081 TCGGCGACGACG 1140
DB 1415 TCGGCGACGACG 1474
1141 CGCTCTTCAAGTGGTCTTACGAGACGCTGCTGACATCGATGAGACTTCCGCGCACCGG 1200
DB 1475 CGCTCTTCAAGTGGTCTTACGAGACGCTGCTGACATCGATGAGACTTCCGCGCACCGG 1534
1201 CAAGCGCAGCAGCGG 1260
DB 1535 CAAGCGCAGCAGCGG 1594
1261 TTGGCGGCGCAGTTGCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
DB 1595 TTGGCGGCGCAGTTGCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1654
1321 AGAGTCATCAGATCAGATCGCAAGGACGAGAGATGCGGAGATTTTCGCGGCACTGG 1380
DB 1655 AGAGTCATCAGATCAGATCGCAAGGACGAGAGATGCGGAGATTTTCGCGGCACTGG 1714
1381 AAGTTTCGCGGCGG 1440
DB 1715 AAGTTTCGCGGCGG 1774
1441 ATCGCGCAGCTAGCGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1488
DB 1775 ATCGCGCAGCTAGCGTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1822

ORGANISM	Heliothis virescens	Db	875	TATGATGGTTATCAGTTGGATCTACAACTACAGATGAAGGGGGCGAGATATAAGCAGT	934
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothis.	Qy	601	TTTGTACGAATCGGAATGGAGTTAATAGGAGTCCCGGCAAGCCACAGATCTAC	660
REFERENCE	1	Db	935	TTTGTACGAATCGGAATGGAGTTAATAGGAGTCCCGGCAAGCCACAGATCTAC	994
AUTHORS	Adamczewski, M.D., Schulte, J.D. and Oellers, N.D.	Qy	661	TACAACTGTGTCCGGAGCCATACATGACATCACGTTTCCGGTGGTGATCCGGAGAAA	720
TITLE	Nucleic acids encoding acetylcholin-receptor subunits from insects	Db	995	TACAACGTGTTCGGAGCCATACATGACATCACGTTTCCGGTGGTGATCCGGAGAAA	1054
JOURNAL	Patent: EP 0962528-A 3 08-DEC-1999;	Qy	721	ACGCTCTACTACTTCTTCAATCTGATCGTCCCTCGGTGCTCATCGCTCCATGGCTCTA	780
FEATURES	BAYER AG (DE)	Db	1055	ACGCTCTACTACTTCTTCAATCTGATCGTCCCTCGGTGCTCATCGCTCCATGGCTCTA	1114
source	Location/Qualifiers	Qy	781	TTGGGGTTTCACTTGCCTCCAGACTCCGGAGAAAAGTTGCTTTAGTGTGACGATATTA	840
	1. 3700	Db	1115	TTGGGGTTTCACTTGCCTCCAGACTCCGGAGAAAAGTTGCTTTAGTGTGACGATATTA	1174
CDS	/organism="Heliothis virescens"	Qy	841	CTGCTGTTGACGGTGTTCCTCAACATGTCGGCGAGACGATCCAGCGAGCTCGACGCC	900
	/mol_type="unassigned DNA"	Db	1175	CTGCTGTTGACGGTGTTCCTCAACATGTCGGCGAGACGATCCAGCGAGCTCGACGCC	1234
	/db_xref="taxon:7102"	Qy	901	GTGCGCTTGTCTCGGACACTTCAACTGCATCATGTTCAATGGTGGCTTCTCCGTGCTC	960
	/note="unnamed protein product"	Db	1235	GTGCGCTTGTCTCGGACACTTCAACTGCATCATGTTCAATGGTGGCTTCTCCGTGCTC	1294
	/codon_start=1	Qy	961	TCACACATGATGCTTCACTTCACTACACACCGGACCGACACACTCAGAAATGAGTAT	1020
	/protein_id="CAC07500.1"	Db	1295	TCCACCATGATGCTTCACTTCACTACACCGGACCGACACACTCAGAAATGAGTAT	1354
	/db_xref="GI:9996845"	Qy	1021	TGATTCGTTGCGGTGTTCTTTATTTGGTCTCGGTGGTGTGCGCATGTCAAGGCCCGGC	1080
	/db_xref="REMBL:CAC07500"	Db	1355	TGATTCGTTGCGGTGTTCTTTATTTGGTCTCGGTGGTGTGCGCATGTCAAGGCCCGGC	1414
	/translation="MGRARRSHLAAPAGLLLLCLLWPRGARGVHEKRLHLHLDH YNYLERPVNESPQLQSLFGLTLMQIIDVDEKQLITNLKLEWMDNLRNTSDF GGVKDLRVPRLHWKPDVLMYNGADEGFDSTPTNVVNRNGSLYVPPGIFKSTCKI DITWFEDDORCEMKFGSTYDGVOLDLQDEGGDISSFVINGEWELIGVPEKRE IYNCCPEPIDITFAVIRKTLXYFNLLI VPCVLTFASMLLGFLLPDPGKELSLG VTLLSLTFLNVAETMPATSDAPELLGTFFNCIMFWASSVVSILLIINLHHRAD THMSDMIRCVFLWLPVLRMRSPGSAITPPPARVPPPPDLERLRSKSLLANVLD IDDFRFPQAQCQCRRYRGGEAGLAHSCFGVDYBELSLILKEIRVITDQMRKD DEDADISRDWKEAMVVDRLCIIFTIITLTIATLAVLLSAPHIMVS"	Qy	1081	TCGGCGACGACGCGCGCGCGCGGTACTCTCCGCGCGGACCTCGGAGCTCGCGAG	1140
ORIGIN		Db	1415	TCGGCGACGACGCGCGCGCGGTACTCTCCGCGCGGACCTCGGAGCTCGCGAG	1474
Query Match	100.0%; Score 1488; DB 6; Length 3700;	Qy	1141	CGCTCTTCCAAGTCTCTTAGGAACTGCTCGACATCGATGACGACTTCCGCAACCG	1200
Best local Similarity	100.0%; Pred. No. 7.2e-291;	Db	1475	CGCTCTTCCAAGTCTCTTAGGAACTGCTCGACATCGATGACGACTTCCGCAACCG	1534
Matches 1488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;		Qy	1201	CAAGCGACGACGCGCAATGCTGCCGATCTACAGGGGGGTGAGGAAATGGCGCGGG	1260
1	ATGGCGGGCGGGCGCGCTCGCACTTGGCGCGCGCGCGCGCTGCTGCTGCTG 60	Db	1535	CAAGCGACGACGCGCAATGCTGCCGATCTACAGGGGGGTGAGGAAATGGCGCGGG	1594
335	ATGGCGGGCGGGCGCGCTCGCACTTGGCGGGCGCGCGCGCTGCTGCTGCTG 394	Qy	1261	TTGCGCGCGACAGTCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1320
61	TGCTGCTGCTGCGCGGGGCGCGCTGCGGTACCAAGAGAGCGGCTACTCCACC 120	Db	1595	TTGCGCGCGACAGTCTTGGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	1654
395	TGCTGCTGCTGCGCGGGGCGCGCTGCGGTACCAAGAGAGCGGCTACTCCACC 454	Qy	1321	AGAGTCATCAGATCAGATCGGCAAGGACGACGACGATTCGGGCGGACTGG	1380
121	CTATTGACCACTACACGTTACTGAGAGGCGCGCTGCTCAACGAGAGCGCGCTG 180	Db	1655	AGAGTCATCAGATCAGATCGGCAAGGACGACGACGATTCGGGCGGACTGG	1714
455	CTATTGACCACTACACGTTACTGAGAGGCGCGCTGCTCAACGAGAGCGCGCTG 514	Qy	1381	AAGTTCGCGCGCATGCTGCGGACAGACTGCTGCTTATTTATCTTACCTGTTCA	1440
181	CTCTCTCTGCGCTCAGCTCATGATCATCGACGTGGAGAGAAACAGCTTTTA 240	Db	1715	AAGTTCGCGCGCATGCTGCGGACAGACTGCTGCTTATTTATCTTACCTGTTCA	1774
515	CTCTCTCTGCGCTCAGCTCATGATCATCGACGTGGAGAGAAACAGCTTTTA 574	Qy	1441	ATCGCCACGCTAGCGGTGCTGCTGCTGCGCGCACACATCATGCTGCTG 1488	
241	ATAACAAACATCTGGCTAAACCTAGTGGAAATGATGAACCTTCAAGTGAAC 300	Db	1775	ATCGCCACGCTAGCGGTGCTGCTGCTGCGCGCACACATCATGCTGCTG 1822	
575	ATAACAAACATCTGGCTAAACCTAGTGGAAATGATGAACCTTCAAGTGAAC 634	LOCUS	ES8347	3701 bp DNA linear PAT 18-JUN-2001	
301	GATTCGCGGGTCAAGATTTAAGTGGCCACCCACAGACTATGGAACACGAGTC 360	DEFINITION	Nucleic acid encoding insect acyl choline receptor subunit.		
635	GATTCGCGGGTCAAGATTTAAGTGGCCACCCACAGACTATGGAACACGAGTC 694	ACCESSION	ES8347		
361	CTTATGTACACAGCGCGAGAGGGTTGACAGCAGCTGATCCAAACGAACTGGTGTG 420	VERSION	ES8347.1	GI:13019346	
695	CTTATGTACACAGCGCGAGAGGGTTGACAGCAGCTGATCCAAACGAACTGGTGTG 754	KEYWORDS	JP 2000023680-A/2.		
421	CGGAACACGCTGCTGCTGCTGCTGCTGCGCGCGGCACTTCAAGAGCACCTGCAAGATC 480	SOURCE	Heliothis virescens (tobacco budworm)		
755	CGGAACACGCTGCTGCTGCTGCTGCTGCTGCGCGCGGCACTTCAAGAGCACCTGCAAGATC 814	ORGANISM	Heliothis virescens		
481	GACATCACTGTTCCCTTTCAGACCAACGATCGAGATGAAGTTTGGCAGCTGGACT 540		Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;		
815	GACATCACTGTTCCCTTTCAGACCAACGATCGAGATGAAGTTTGGCAGCTGGACT 874				
541	TATGATGGTTATCAGTTCGATCTACAACTACAGGATGAAGGGGGCGGAGATATAAGCAGT 600				

Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia; Noctuoidea; Noctuidae; Heliothinae; Heliothis.	
REFERENCE	1. (bases 1 to 3701)
AUTHORS	Martin,A., Nadja,E. and Thomas,S.
TITLE	Nucleic acid encoding insect acetyl choline receptor subunit
JOURNAL	Patent: JP 2000023680-A 2 25-JAN-2000; BAYER AG
COMMENT	
OS	Heliothis virescens
PN	JP 2000023680-A/2
PD	25-JAN-2000
PF	26-APR-1999 JP 1999118159
PR	04-MAY-1998 DE 19819829.9
PI	MARTIN ADAMUTSUJESUKI,NADJA ERASU,THOMAS SCHULTE PC
C1	C12N15/09,A01K67/033,C07K14/705,C07K16/28,C12N1/21,C12N5/10, PC
C1201/68.	
PC	G01N33/15,G01N33/50/(C12N1/21,C12R1/19),C12N15/00,C12N5/00 CC
FEATURES	
source	
ORIGIN	
Query Match 100.0%; Score 1488; DB 6; Length 3701; Best Local Similarity 100.0%; Pred. No. 7.2e-291; Mismatches 0; Indels 0; Gaps 0; Matches 1488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1 ATGGCGGGCGGGCGCGCTCGCATTGGCGGGCGCGCGGGCGCTGCTGCTGCTG 60
Db	335 ATGGCGGGCGGGCGCGCTCGCATTGGCGGGCGCGCGGGCGCTGCTGCTGCTG 394
QY	61 TGCTGCTGCTGCGGAGGGGGGCGCTGCGGGGTACACGAGAGCGGCTACGCAAC 120
Db	395 TGCTGCTGCTGCGGAGGGGGGCGCTGCGGGGTACACGAGAGCGGCTACGCAAC 454
QY	121 CTATTGGACCACTCAACAGTACTGAGAGGCGCGTCTGCAAGAGAGAGCGCGTGCAG 180
Db	455 CTATTGGACCACTCAACAGTACTGAGAGGCGCGTCTGCAAGAGAGAGCGCGTGCAG 514
QY	181 CTCTCTTGGCTCAAGCTCATGCGATCATGCGTGGAGAGAGAGAGAGAGAGAGTTTA 240
Db	515 CTCTCTTGGCTCAAGCTCATGCGATCATGCGTGGAGAGAGAGAGAGAGAGAGTTTA 574
QY	241 ATACAAACATCTGGCTAAACTAGAGTGGATGATGATGATGATGATGATGATGATGAT 300
Db	575 ATACAAACATCTGGCTAAACTAGAGTGGATGATGATGATGATGATGATGATGATGAT 634
QY	301 GATTTGGCGGGGTCAAGATTTAAGAGTGCACCCACAGATATGGAACACAGAGCTC 360
Db	635 GATTTGGCGGGGTCAAGATTTAAGAGTGCACCCACAGATATGGAACACAGAGCTC 694
QY	361 CTTATGTACACAGCGGAGAGAGGTTTCAAGAGAGAGTATCAACAGAGAGAGTGGTG 420
Db	695 CTTATGTACACAGCGGAGAGAGGTTTCAAGAGAGAGTATCAACAGAGAGAGTGGTG 754
QY	421 CGGAACACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
Db	755 CGGAACACAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 814
QY	481 GACATCACTGGTTTCCCTTTCGACGACCAACAGATGCGAGATGAAGTTTGGCAGCTG 540
Db	815 GACATCACTGGTTTCCCTTTCGACGACCAACAGATGCGAGATGAAGTTTGGCAGCTG 874
QY	541 TATGATGTTATCAGTTGGATCTCAACTACAGATGAAGGGGGCGGAGATATTAAGCAGT 600
Db	875 TATGATGTTATCAGTTGGATCTCAACTACAGATGAAGGGGGCGGAGATATTAAGCAGT 934
QY	601 TTTTGTACGAATGGGAATGGAGTTAATAGAGTCCCGGCAAGCGCAAGCATCTAC 660
Db	935 TTTTGTACGAATGGGAATGGAGTTAATAGAGTCCCGGCAAGCGCAAGCATCTAC 994

QY	661 TACAACTGTTTCGGAGGCATACATGACATCATGTTTTCGGTGGTGTATTCGGAGGAAA 720
Db	995 TACAACTGTTTCGGAGGCATACATGACATCATGTTTTCGGTGGTGTATTCGGAGGAAA 1054
QY	721 AGCTCTACTACTCTTCAATCTGATCGTGGCCCTCGTGTCTCATCCCTCCATGCTCTA 780
Db	1055 AGCTCTACTACTCTTCAATCTGATCGTGGCCCTCGTGTCTCATCCCTCCATGCTCTA 1114
QY	781 TTGGGGTTTCACTTTCCTCCAGACTCCGGAGAAAAAGTTGTCTTTAGGTGTGACGATATTA 840
Db	1115 TTGGGGTTTCACTTTCCTCCAGACTCCGGAGAAAAAGTTGTCTTTAGGTGTGACGATATTA 1174
QY	841 CTGTCTGTGACGGTGTCTCTCAACATGTTGGCGGAGACGATGCCAGACGCTCGACGCC 900
Db	1175 CTGTCTGTGACGGTGTCTCTCAACATGTTGGCGGAGACGATGCCAGACGCTCGACGCC 1234
QY	901 GTGCCCTTGTCTCGGCACCTACTTCAACTGCACTATGTTTCATGGTGGTTCCTCCGTGCTC 960
Db	1235 GTGCCCTTGTCTCGGCACCTACTTCAACTGCACTATGTTTCATGGTGGTTCCTCCGTGCTC 1294
QY	961 TCCACCACTACTGATCTCTCAACTACCAACCGGCAACGACACTCAAGAAATGATGAT 1020
Db	1295 TCCACCACTACTGATCTCTCAACTACCAACCGGCAACGACACTCAAGAAATGATGAT 1354
QY	1021 TGGATTGTTGGTGTCTCTTATTTATGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 1080
Db	1355 TGGATTGTTGGTGTCTCTTATTTATGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 1414
QY	1081 TCGCGCAGCAGCGCGCGCGCGCTACTCTCGCGCGCGCGCGCTGAGCTGCGCGAG 1140
Db	1415 TCGCGCAGCAGCGCGCGCGCGCTACTCTCGCGCGCGCGCGCTGAGCTGCGCGAG 1474
QY	1141 CGCTCTTCAAAGTCTCTTAGCGAACTGCTCGACATCGATGACGATTCGCGCACCGG 1200
Db	1475 CGCTCTTCAAAGTCTCTTAGCGAACTGCTCGACATCGATGACGATTCGCGCACCGG 1534
QY	1201 CAAGCGAGCAGCGCATGCTGCGATCTACAGGGGGGTGAGGAGAAATGGCGCGGG 1260
Db	1535 CAAGCGAGCAGCGCATGCTGCGATCTACAGGGGGGTGAGGAGAAATGGCGCGGG 1594
QY	1261 TTGGCGGCGCAGCTTGTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1320
Db	1595 TTGGCGGCGCAGCTTGTCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1654
QY	1321 AGAGTCATCAGATCAGATCGGCAAGAGACGAGATGCGGACATTTTCGCGGAGCTGG 1380
Db	1655 AGAGTCATCAGATCAGATCGGCAAGAGACGAGATGCGGACATTTTCGCGGAGCTGG 1714
QY	1381 AAGTTCCGCGCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1440
Db	1715 AAGTTCCGCGCATGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1774
QY	1441 ATGCCACGCTAGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1488
Db	1775 ATGCCACGCTAGCCGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1822

RESULT 4	AF321446	2023 bp	mRNA	linear	INV 29-APR-2002
LOCUS	AF321446				
DEFINITION	Drosophila melanogaster nicotinic acetylcholine receptor Dalphaf				
	subunit variant type II (mAcRalpha-30D) mRNA, complete cds,				
	alternatively spliced.				
ACCESSION	AF321446	GI:20152846			
VERSION	AF321446.1				
KEYWORDS					
SOURCE	Drosophila melanogaster (fruit fly)				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
	Ephydroidea; Drosophilidae; Drosophila.				
REFERENCE	1 (bases 1 to 2023)				

Db 472 GTACGTGCCCCCTGGTATCTTCAAGAGCACATGCAAGATGACATCACGTGTTCCCAT 531
QY 501 CGACGACCAACGATCGAGATGAAGTTTGGCAGCTGGACTATGATGGTTATCATGTTGA 560
Db 532 TGATGACCAACATTCGGAATCGAAATTCGGTAGTTGGACTTACGATGGAATTCAGTTGA 591
QY 561 TCTACAACTACAGGATGAAGGGCGGAGATATAGCAGTTTGTGTCAGGAATGGCGAATG 620
Db 592 TTTGGTTTGTAGTTCCGAGATGGAGGGAICTTTCCGANTTTCATAAATATGGCGAGTG 651
QY 621 GGAGTTAATAGAGTCCCGGCAAGCGCAACGAGATCTACTACAACTGTGTGTCGGAGCC 680
Db 652 GTACTTGTGTCATGCGCGGAAAGAAATACGATAGTCTACGCGTCTGCGCCAGAACC 711
QY 681 ATACATCCACATCAGTTTGGGTGTGATCGGAGGAAACGCTCTACTACTTCTTCAA 740
Db 712 ATATGTCATATCACCTTTACTATACAAATTCGTGCGCGTACATATATATTTTCAA 771
QY 741 TCTGATCGTGCCTCGCTCATCGCTCCATGCTCTATTGGGGTTTCACCTTGCCTCC 800
Db 772 TTTAATTGCAATGTGTCTAATCTCATCGATGCCCTACTGGGCTTCACATTCGCGCC 831
QY 801 AGACTCCGAGAAAAGTTGTCTTTAGTGTGACGATATTAATGCTGTGACGGTGTCT 860
Db 832 GGATTCGGCGGAGAACTGACGCTGGCGTAACATACTACTATCATTAACAGTATTCT 891
QY 861 CAACATGTGTGGGAGAGATGCCAGCGACGCTCGACGCGCTGCTCGGACCTA 920
Db 892 AAACCTTGTGCGCGAGTCCATGCGACAAACGTCGATGCTGTCTCTATAGGCACCTA 951
QY 921 CTTCAACTGCATCATGTTTCATGTTGGTCTCTCCCTGCTCTCCACATCACTGATCCTCAA 980
Db 952 CTTCAATTGCATCATGTTTCATGTCGCTGCTGCTGCTGCTGCTGCTGCTGCTCAA 1011
QY 981 CTACCAACCGGCAACGAGACACTCAGCAATGAGTGAATGATGTTGTTGTTGTTCT 1040
Db 1012 CTACCACCATGCGACAGCGGACATTCAGGAGATGCCACCGTGATCAAGTCCGTTTCT 1071
QY 1041 TTATTGGTGGTGGTGTGCTGCGATGCTCAGCGCGCGGCTCGGCGAGCGCGCGCC 1100
Db 1072 ACAATGGTGGCTGGATCTTTCGAATGGGTGCGACCGCGTCCGAAGATTACACGCAAAAC 1131
QY 1101 GCGCGCGCTA---CCTCGCGCGCGGAGCTGGAGCTGGCGAGCGCTCTCCCAAGTCGCT 1157
Db 1132 AATACTATTAGCAATCGCATGAAGAGCTGGAGCTAAGGAGCGCTCTCCAATCCCT 1191
QY 1158 CTTAGCGAAGTGTGCGATGATGACGACTTCGCGCACCGCGAAGCGAGCGACGCGGA 1217
Db 1192 GGTGCGCAATGTCTCGACATGCGACGAGACTTTCGCGCACCAATATCTGGCTCCCAAAC 1251
QY 1218 ATGCTGCCGATACTACAGGGGGGTGAGAGAATGCGCGGGTGGCGGCGCAGTTG 1277
Db 1252 CGCCATTGGCTGTCGCGCAGCTTCGTCGCGGCCACACAGGTGGAGGAGCATCACACGGC 1311
QY 1278 CTTGGGTGCG---ACTACGAGCTCTCCCTCATCTGAAGGAGATTAGAGTCAATCACAGA 1334
Db 1312 CATCGGCTGCAATCACAAAGATCTTCATCTAATCTCAAGAAATTCGCAATTTATACGCG 1371
QY 1335 TCAGATGCGCAAGGACGAGAGATGCGGACATTCGCGCGGAGCTGGAAGTTGCGCGCCAT 1394
Db 1372 GCGGATGCGCAAGCTGACGAGCGGAAATGATCGCGGATTTGATCGCGGATTCGCGGCAAT 1431
QY 1395 GGTCTGTGACAGCTGTGCTTATTATCTTTTACCTGTTTCAAAATCATCGCCAGCCTAGC 1454
Db 1432 GGTGTGGATGATTTTGTTAATGTTTTCACGCTCTTCAGATATATGCAAGGTTAC 1491
QY 1455 GGTGTGTGTCGCGCGCACACATCATGTTG 1485
Db 1492 GGTGTGTCTCGCTCCGACATATATGTTG 1522

RESULT 7

AF321447 2023 bp mRNA linear INV 29-APR-2002
LOCUS Drosophila melanogaster nicotinic acetylcholine receptor Dalph6
DEFINITION subunit variant type III (nAcRalpha-30D) mRNA, complete cds,
alternatively spliced.
ACCESSION AF321447
VERSION AF321447.1 GI:20152848
KEYWORDS Drosophila melanogaster (fruit fly)
SOURCE Drosophila melanogaster
ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 2023)
AUTHORS Grauso, M., Reenan, R.A., Culetto, E. and Sattelle, D.B.
TITLE Novel Putative Nicotinic Acetylcholine Receptor Subunit Genes,
Dalph6, Dalph6 and Dalph7, in Drosophila melanogaster Identify a
New and Highly Conserved Target of Adenosine Deaminase Acting on
RNA-Mediated A-to-I Pre-mRNA Editing
JOURNAL Genetics 160 (4), 1519-1533 (2002)
MEDLINE 21969411
PubMed 11973307
REFERENCE 2 (bases 1 to 2023)
AUTHORS Grauso, M. and Sattelle, D.B.
TITLE Direct Submission
JOURNAL Submitted (15-NOV-2000) Human Anatomy and Genetics-Functional
Genetics Unit, University of Oxford, South Park Road, Oxford OX1
3QX, UK
FEATURES
source Location/Qualifiers
1..2023 /organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/chromosome="2"
/map="30D1"
/dev_stage="embryo"
1..2023 /gene="nAcRalpha-30D"
379..1863 /gene="nAcRalpha-30D"
/note="ion channel; neurotransmitter transmembrane
receptor; alternatively spliced; contains exons 3b and 8b"
/codon_start=1
/product="nicotinic acetylcholine receptor Dalph6 subunit
variant type III"
/protein_id="AA013394.1"
/db_xref="GI:20152849"
/translation="MDSPLPASLSFVLLIFLAIKESCQGPHEKRLNHLSTVNTL
ERPVAESEPLEVEFGLTQIIDVEKNQLLI TNLMSLEWNYLNRNTEYGGVK
DLRITPNKLPDVMYNSADEGDTYHTNIVVKGSGCLYVPPGIFKSTCKMDITW
PFDDQHCMEKFGSWYDGNLDLVNSEDGDSDFITNGEWYLLAMPKNTIYA
CCBPYVDITFTIOIRRTLYFFENLIVPCVLISSMALLGFTLPDSEKLTIGVTIL
LSLVFLNVAETLPQVSDAIPILGTYFNCIMFVASSVLTIVVLYNHHRTADHEM
PFWKSVLQWLPWLMRGPRKTRITLLSNRKELEKRSKSLANVLDD
DFRHTISGOSTAIGSSAFGRPTTVEHTTAIGCNHKLHLILKELQFITARMRKADD
BAELIGDWKFAAMVDFRLIVFTLTITATVTLLSAPHIVQ"
794
variation /gene="nAcRalpha-30D"
/note="results in glycine to serine substitution; compared
to the sequence deposited in GenBank Accession Number
AE003626"
846 /replace="a"
variation /gene="nAcRalpha-30D"
/note="results in asparagine to serine substitution;
compared to the sequence deposited in GenBank Accession
Number AE003626"
846 /replace="a"
ORIGIN
Query March 41.1%; Score 611.8; DB 3; Length 2023;
Best Local Similarity 64.4%; Pred. No. 1.6e-113;
Matches 948; Conservative 0; Mismatches 517; Indels 6; Gaps 2;

Db 1753 GGTGTGCCGGCAACGTAACGAGATCTATTACAACTGCTGCCGGAACCCCTATTATAGAC 1812
QY 691 ATCAGTTTGGGTGGTGGTATCGGAGGAAGCGTCTACTACTTCTCAACTGATGCTG 750
Db 1813 ATCAGTTTGGGTGGTGGTATCGGAGGAAGCGTCTACTACTTCTCAACTGATGCTG 1872
QY 751 CCCTGCTGCTCATCGCTCCATGCTCTATTGGGGTTCACCTTCCCTCCAGACTCCCGGA 810
Db 1873 CCTTGTGACTGATTCCTCCATGCTTGCAGATTCACCTTCCCTCCAGACTCCCGGT 1932
QY 811 GAAAGTTGCTTTAGGTGTGACGATATTACTGCTGTGACGGTGTCTCTCAACATGCTG 870
Db 1933 GAAAGTTGCTTTAGGTGTGACGATATTACTGCTGTGACGGTGTCTCTCAACATGCTG 1992
QY 871 GCGGAGAGATGCGGAGGAGCGTCCGACGCGGCTTGTCTGCGGACCTACTTCAACTGC 930
Db 1993 GCGGAGAGATGCGGAGGAGCGTCCGACGCGGCTTGTCTGCGGACCTACTTCAACTGC 2052
QY 931 ATCATGTTTCATGGTGGTTCCTCCGTGCTCTCCACCATGATCTCAACTTACCACAC 990
Db 2053 ATATGTTTATGGTGGTTCCTCCGTGCTCTCCACCATGATCTCAACTTACCACAC 2112
QY 991 CGGACGAGACACTCAGAAATGATGATGATGATGATGATGATGATGATGATGATGATG 1050
Db 2113 CGAAATGCTGATACGACGAAATGTCGAAATGATGATGATGATGATGATGATGATG 2172
QY 1051 CGGTGGTGTGCGCATGTACGCGCGCG- - - - -CTCGCGGAGACGCGCGC 1097
Db 2173 CATGATGATGCGAATGATGCGCCAGGACGACGCGTGTCTGATGATGCGGCGCGC 2232
QY 1098 GCGGCGCGGCTACCTCCGCGC- - - - -CCGACCTGAGCTGGC 1137
Db 2233 CCCTGTTCCGACACATCTCCGAGCGGAAGACAGATCTCTCGACGTTGAGCTGAA 2292
QY 1138 GAGCGCTCTCAAGTGTCTTACGAAAGTGTCTGAGATGATGATGATGATGATGATG 1197
Db 2293 GAGCGCTCTCAAGTGTCTTACGAAAGTGTCTGAGATGATGATGATGATGATGATG 2352
QY 1198 - - - - -CCGAGCGCAGCAGCGCAATGCTCGCATACTACGAGCGGG 1241
Db 2353 AATGTGCGCCATGACGCGCGGGAACATGCGCACACACCGCGCTTCTATCGCAG 2412
QY 1242 TGAGGAGATGCGCGGGTGTGGCGGCGACAGTTGCTTGGTG- - - - - 1285
Db 2413 GTTATGACAAAGCGAGTGGCAGCATGGGCAATGGCAGCACCGCAATGCGCAT 2472
QY 1286 - - - - -TCGACTAGCAGCTCTCCCTCATTTCTG 1311
Db 2473 GCGGTACCCATCATACGTGATCAAAATCATCACTGAATATGAATAGTTTAACTTA 2532
QY 1312 AAGGAGATTAGTTCATCAGATCAGATGCGCAAGGAGCAGCAAGATGCGGACATTTTCG 1371
Db 2533 AAGGAAATTCGTTTATTAATCATGATGATGATGATGATGATGATGATGATGATGATG 2592
QY 1372 CGGACTGGAATGCGCGCATGCTGTGAGACAGCTGTGCTTATTTATTTTACCCTG 1431
Db 2593 AATGATTGGAATTTGACGATGCTGTGAGACAGCTGTGCTTATTTATTTTACAATG 2652
QY 1432 TTCGAATCATGCGCAGCTAGCGGTGCTGTGCGGCGCACACATCATGCTGTGCTG 1488
Db 2653 TTCGAATCATGCGCAGCTAGCGGTGCTGTGCGGCGCACACATCATGCTGTGCTG 2709

RESULT 10
AF143847

LOCUS AF143847 3029 bp mRNA linear INV 27-MAY-1999
DEFINITION Heliothis virescens putative nicotinic acetylcholine receptor alpha
7-2 subunit mRNA, complete cds.

ACCESSION AF143847

VERSION AF143847.1 GI:4895006

KEYWORDS

SOURCE Heliothis virescens (tobacco budworm)

ORGANISM Heliothis virescens

Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
Noctuoidea; Noctuidae; Heliothinae; Heliothis.
1 (bases 1 to 3029)
Schulze, T., Oellers, N. and Adamczewski, M.
Putative alpha subunits of insect nicotinic acetylcholine receptors
more similar to vertebrate alpha 7 subunits and C. elegans Ce21
than to other insect nicotinic acetylcholine receptor alpha
subunits

JOURNAL

Unpublished

2 (bases 1 to 3029)

Schulze, T., Oellers, N. and Adamczewski, M.

Direct Submission

Submitted (19-APR-1999) ZP-BTB, Bayer AG, Bldg. Q 18., Leverkusen

51368, Germany

Location/Qualifiers

1. 3029

/organism="Heliothis virescens"

/mol_type="mRNA"

/db_xref="taxon:7102"

95-1600

/note="hvnachra7-2"

/codon_start=1

/product="putative nicotinic acetylcholine receptor alpha

7-2 subunit"

/protein_id="AAD32698.1"

/db_xref="GI:4895007"

/translation="MAPLAALALLLPVSEQGHKRLNALLANNVTLERPVANE

SEPLEVRFGLTQIIDVDKQQLITNWLSEWLDYNLRKNSDSEYGGVXDLRITPN

KLWPDVLYNSADEGFDGTQTVVVRSGSCLYVPGIKAWPFDDQHQ

CDKMFQGWTDVQLNDLVLDKDEAGDLSDFITNGEWLIGMFKKNTITYACPEPYV

DVTFTIMRRRLTYFNNLIVPCVLISSMLLGLTLPDSEKLTGLVTILLSTVFL

NLVAETLPOVSDAIPLLGTENCIMEMVASSVLTVVVLYNHHRTADTHEMPOIKSV

FLQWLPILNRSRFGKIKTRKIMNTRMRELEKRSKSLIANLVLDIDDPHGGP

PPNSTASTNGLPGCSIFRTDFRRSFVRPSTMEDVGGGLGSHHRLHLILRELQPIA

RMKKADEAEELISDWKFAAMVVDVRFCLFVFTLTIIATVAVLLSAPHITVQ"

ORIGIN

Query Match 38.6%; Score 575; DB 3; Length 3029;
Best Local Similarity 64.6%; Pred. No. 4.6e-106;
Matches 929; Conservative 0; Mismatches 460; Indels 48; Gaps 3;

QY 97 CACGAGAGCGGCTACTGCACCCCTATTGGACCCTACAACTACTGAGAGGCGCCGTC 156
Db 158 CACGAGAGAGACTCCTGAACGCGTGTCTGCGAACTACAAACCCCTGAGCGACCGGTG 217
QY 157 GTCAACGAGAGCGACCGCTGAGCTCTCTTGGCCTCAGCTCATGCTCATGATCATGAC 216
Db 218 GCCAACGAGAGCGAACCGCTAGAGGTTCAGGTTCGGCTTGCCTTGCACGAAATCATTCAC 277
QY 217 GTGACAGAGAGAACCGAGCTTTTAAACAAACATCTCGCTAAACTAGAGTGAATGAT 276
Db 278 GTGACAGAGAGATCACTACTTATTAACCAATATATGGCTGTCTGTTGAGTGAATGAC 337
QY 277 ATGAACCTTGGGTGAACACTTCAGATTTCGCGGGGTCAAAGATTTTAAGAGTGCACCC 336
Db 338 TACAACCTGAGGTGAACGAGCGAGTATGCGGGGTCAAAGACCTCAGGATCACGCC 397
QY 337 CACGACTATGGAACGAGAGCTCTTATGTACACGCGGAGCAAGAGGTTCACAGC 396
Db 398 AACAGTTGTGGAACCGGAGCTCTTATGTATATAGTGTGAGAGGGTTTTGACGGG 457
QY 397 ACCTATCCAAACGAGCTGGTGTGCGGAACAAACGGCTCGTCTGTAGTGTGCGCGCGC 456
Db 458 ACCTACCAGACCAAGTGGTGTGAGAGCGCGGCGAGTTCCTGTAGTGTGCGACCTGCG 517
QY 457 ATCTTCAAGAGACCTGCAAGATCGACATCATCTGTTTCCCTTCGAGCAGCAACAGTGC 516
Db 518 ATATTCAAGAGACATGCAAGATGGACATCGCGTGTTCCTTCGAGCAGCAACACTGT 577
QY 517 GAGATGAGTTTGGCAGCTGACTTATGATGTTATGATGTTGATCTCAACTACAGGAT 576
Db 578 GATATGAAGTTTCGGTAGCTGGACATATGACGGCAATAGTTGGATCTGGTGTCTAAAGAT 637

```
QY 577 GAAGGGGCGGAGATATACAGAGTTTGTACGAAATGCGGAATCGGAGTTAATAGGATC 636
D 638 GAGCAGCGCGGATCTATCGGACTTATACAAATGGGAGTGTATCTAATAGGAATG 697
QY 637 CCGGGAAGCGCAAGAGATCTACTACAACTGTGTCCGAGGCAATACATCGACATCAG 696
D 698 CCAGGCAAAAGAACAAATACATACGATGCGTGTCCCGAGCCCTAGTGAGAGTCAAC 757
QY 697 TTTGCGGTGGTATCCGGAGGAAACGCTCTACTACTTCTTCAATCTGATCGTCCCTGC 756
D 758 TTCACCATCATGATAAGAGAGCAACCTTGCTACTTCTTCAACCTGATGCTCCGTC 817
QY 757 GTGCTCATCGCTTCATCGGCTCTATTTGGGTTACCTTGCTCCAGACTCCGAGAAAG 816
D 818 GTGCTGATCTCATCGATGGCACTCTCGGCTTCACTGCGCAACAGACTCCGAGAGAAA 877
QY 817 TTGTCCTTTAGGTGACGATATTAATCTGCTGTGACGGTGTTCCTCAACATGCTGGCGAG 876
D 878 CTACACTTGGAGTCACTATCTTCTATCGCTGACGGTGTTCCTCAACCTGCTAGCCGAG 937
QY 877 ACGATCGACGAGCTCGGAGCGCGTGCCTTGTCTCGGCACTACTTCAACTGCATCATG 936
D 938 ACCCTGCCACAGTCTCCGAGCGCTATCCCTGTTAGGAGCTACTTCAATTGCATCATG 997
QY 937 TTCATGTTGGCTTCTCTCGTGTCTCCACCATACTGATCTCACTCACTACCAACCGGCAC 996
D 998 TTCATGTTAGCTGTCTGTGTTACTGACTGTGTGTTACTCAATTACCACTATCGAACA 1057
QY 997 GCAGACTCAAGAAATGAGTATGATGTTGCTGTGCTGTTCTTATTTGGCTCCGCTGG 1056
D 1058 GCTGATATACATGAATGCCACAGTGTGATAAATCAGTATTCCTACATGTTGCCATGG 1117
QY 1057 GTGCTGCGCATGTCACGCGCGCTCGGCG---ACGACGCGCGCGCGCGCGCGGTACCT 1113
D 1118 ATACTGCGAATGTGAGCGCCAGGAGAAAGATACCAGGAAGACTATAATGATGAACAG 1177
QY 1114 CCGCGCGCGGACTGAGAGCTGCGGAGCGTCTCCAGAGTCTCTTACAGTCTCTAGCGAA 1173
D 1178 AGGATGAGGAGCTGGAACCTGAAGGAGGTCGTGCGAAGTCTTGTGCGGAAATGTTCTA 1237
QY 1174 GACATGATGACGACTTCCGCGCACCGCAAGCGCAGC-----AGCGCAATCTGCGG 1226
D 1238 GATATTGATGATGACTTCAGACAGCGGCTCCGCTCTTAACAGTACTGCTCGACCGGG 1297
QY 1227 ATACTACAGGGGGTGAAGAGATGCGCGGGTGTGCGCGCACAGTTGCTTCG---- 1282
D 1298 AATTGGGACCTGGGTGCTCAATATTCCGACAGGATTTCCGTCGGTGTGCTCGCTCGG 1357
QY 1283 -----GTGTCGACTAGGACTCTCCCTCATTT 1308
D 1358 TCCACGATGGAAGACGTGGGCGCGGGTGGGTAGCCACCACTCGCGAGCTGCACCTCATA 1417
QY 1309 CTGAAGGAGATTAGAGTATCACAGATCAGATGCGCAAGGACGACGAGATGCGGACATT 1368
D 1418 CTGAGAGGCTGCAAGTTCATCCGCGCAGATGAGAGGCTGATGAGGAAGCGAGCTG 1477
QY 1369 TCGCGCACTGGAAGTTGCGCGCATGTTGCTGAGACAGACTGTGCTTATTTTAC 1428
D 1478 ATCAGGACTGGAAGTTGCTGCGATGTTGCTGATAGGTTTCCCTGTTGCTGTTCACA 1537
QY 1429 CTGTTCAATCATCGCACGCTAGCTGCTGTGCTGCTGCGGCGCACATCATGTTG 1485
D 1538 CTTTTCATATCATCGGACAGTACTGTCTGTTATCGGCAACCGCATATCATGTT 1594

RESULT 11
E58348 58348 3109 bp DNA linear PAT 18-JUN-2001
LOCUS
DEFINITION Nucleic acid encoding insect actyl choline receptor subunit.
ACCESSION E58348
VERSION E58348.1 GI:13019347
KEYWORDS JP 2000023680-A/3.
```

```
SOURCE Heliothis virescens (tobacco budworm)
ORGANISM Heliothis virescens
          Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
          Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
          Noctuoidea; Noctuidae; Heliothinae; Heliothis.
REFERENCE 1 (bases 1 to 3109)
AUTHORS Martin,A., Nadja,E. and Thomas,S.
TITLE Nucleic acid encoding insect actyl choline receptor subunit
JOURNAL Patent: JP 2000023680-A 3 25-JAN-2000;
          BAYER AG
COMMENT OS Heliothis virescens
          PN JP 2000023680-A/3
          PD 25-JAN-2000
          PF 26-APR-1999 JP 1999118159
          PR 04-MAY-1998 DE 19819829.9
          PI MARTIN ADAMUTSUEUSUKI,NADJA ERASU,THOMAS SCHULTE PC
          C12N15/09,A01K67/033,C07K16/705,C07K14/28,C12N1/21,C12N5/10, PC
          C12Q1/68.
          PC G01N33/15,G01N33/50//(C12N1/21,C12R1:19),C12N15/00,C12N5/00 CC

FH Key Location/Qualifiers
FT CDS Location/Qualifiers
FEATURES             Location/Qualifiers
     source            1..3109
                        /organism="Heliothis virescens"
                        /mol_type="genomic DNA"
                        /db_xref="taxon:7102"
ORIGIN
Query Match      38.6%; Score 575; DB 6; Length 3109;
Best Local Similarity 64.6%; Pred. No. 4.6e-106;
Matches 929; Conservative 0; Mismatches 460; Indels 48; Gaps 3;
QY 97 CACGAGAAGCGGTACTGTCACACCTATTGGACCACTACAACTGCTGGAGAGCCCGTGC 156
D 158 CACGAGAAGAGACTCTCTGAACGCGTGTCTGCGCAACTACAACTGCTGGAGAGCCCGTGC 217
QY 157 GTCAACGAGAGAGCCCGCTGCGAGCTCTCTTGGGCTCAGCTCATGCGAGATCATCGAC 216
D 218 GCCAACGAGAGCAACCGCTAGAGGTGAGGTTCGGCTTGACCTTTCGACCAATCATTTGAC 277
QY 217 GTGACGAGAGAACACAGCTTTTAAACAAACATCTGGCTTAAACTAGAGTGAATGAT 276
D 278 GTGACGAGAGAAATCAACTACTTATTAACCAATATATGGCTGTCTGTTGGAGTGAATGAC 337
QY 277 ATGAACCTTGAGGTGGAACTTCAGATTTTCGGCGGGTCAAAGATTTTAAGAGTGCACCC 336
D 338 TACAACCTGAGGTGGAACGACGAGCTATGGCGGGTCAAGGACCTCAGGATCACGCC 397
QY 337 CACGACTATGGAACACGAGCTCTTATGTACAAAGCGGAGCGAGGTTTCGACAGC 396
D 398 AACAGTTGTGGAAGCGGAGCTCTTATGTATTAATAGTCTGACGAGGTTTTCGACGG 457
QY 397 ACGTATCAACGACGAGTGTGTCGGAACAAACGGCTCGTGTCTGTCTGACGTCGCCCGGC 456
D 458 ACCTACGAGACCAACGTTGTTGTCAGAGCGCGGAGTTCCTGTACGTGCCACTTGGC 517
QY 457 ATCTTCAAGAGCACTTCAAGATCGACATCACTTGGTTCCTTTCGACGACCAACGATGC 516
D 518 ATATTCAAGAGCATCAAGATGGACATCGGTGGTTTCCCTTTCGACGACCAACACTGT 577
QY 517 GAGTGAAGTTTGGCAGCTTATGATGTTTATGATGTTGATCTTCACTACTACACTACAGAT 576
D 578 GATATGAAGTTCGGTAGCTGGACATATGACGGCAATCAGTTGGATCTGGTCTAAAAGAT 637
QY 577 GAAGGGGCGGAGATATAAGCAGTTTGTCAAGTCCGGAATGGGAGTTAATAGGAGTC 636
D 638 GAGCAGCGCGGATCTATCGGACTTCAATCAATGGGAGTGTATCTAATAGGATG 697
QY 637 CCGGCAAGCGCAACGAGATCTACTACAACTGTTTCCGAGGCAATACATCGACATCAG 696
D 698 CCAGGCAAAAGAACACAAATACATACGCTGTGCTGCGGCGGCTACGTTGGAGCTCAC 757
```

697 TTTCGGTGGTATCCGAGGAAACGCTCTACTACTTCTTCAATCTGATCGCCCTGC 756
 758 TTCCATCATGATGAAGAAGACGAACCTTGTAATCTTCTTCAACCTGATCGTCCGTCG 817
 757 GTGCTCATCGCTCATGCTCTATTTGGGTTTCACTTGGCTTCAAGCTCCGAGAAAG 816
 818 GTGCTGATCTCATCGATGGCACTCTCGGCTTCACTGCCACCAAGCTCCGAGAGAAA 877
 817 TTGCTCTTATAGTGTGACGATATTAATCTGCTGTGAGGCTGTTCTCAACATGCTGGCGAG 876
 878 CTCACACTTGGAGTCACTATTTCTTATGCTGACGGTGTCTTCAACCTGTGACCGAG 937
 877 ACGATGCCAGCACTCGGACCGCTGCTTCTGCTCGGACCTACTTCAACTGCACTATG 936
 938 ACCCTGCCACAGTCTCCGACGCTATCCCTCTGTTAGGAGCTACTTCAATTTGCACTATG 997
 937 TTTCATGGTGGCTTCTCTCGTCTCCACCATATGATCTTCAACTGCACTATG 996
 998 TTTCATGGTGGCTTCTCTCGTCTCCACCATATGATCTTCAACTGCACTATG 1057
 997 GCAGACACTCAGAAATGAGTGAATGCTGCTGCTGTTCTTATTTGGCTGCGCTGG 1056
 1058 GCTGATATACATGAATGCCACAGTGGATAAATCAGTATTTCTTCAATGTTGCCATGG 1117
 1057 GTGCTGCGCATGTACAGGCGCGCTGCTGCGG---ACGACGCGCGCGCGCGCTACCT 1113
 1118 ATACTGCAATGTGAGGCGGAGGAAAGATCACAGGAGACTATAATGATGAACAG 1177
 1114 CGCGCGCGGAGCTGGAGCTGCGGAGCGCTCTTCAAGTCTGCTTACGCAAGCTGCTC 1173
 1178 AGGATGAGGAGCTGGAACTGAAGGAGAGGCTGCTGCAAGTCTTCTGCGCAATGTTCTA 1237
 1174 GACATCGATGACGACTTCCGCGCACCGCGAGCGAGC-----AGCGCAATGCTGCG 1226
 1238 GATATGATGATGACTTTCAGACAGCGCCCTCGGCTCTTACAGTACTGCTCGACCGG 1297
 1227 ATACTAGGCGGGGTGAGGAGATGCGCGGGGTGCGCGGCGCAAGTCTGCTCG----- 1282
 1298 AATTGGGACCTGGGTGCTCAATATTCGCGACGGAATTCCTGCGGTGCTTCTGCTGCTCG 1357
 1283 -----GTGCTGACTAGGACTTCTCCTCTCAT 1308
 1358 TCCAGATGGAAGAGCTGGGCGGGGTGGGTAGCCACCATCGGAGCTGCACCTCAT 1417
 1309 CTGAAGGAGATTAGAGTATCACAGATCAGATGCGCAAGGACGAGCAAGATGCGGACATT 1368
 1418 CTGAGAGAGCTGCATTCATCAGCGCCAGATGAAGAGGCTGATGAGGAGCGGAGCTG 1477
 1369 TCGCGGAGCTGGAAGTTCGCGGCAATGCTGCTGAGAGAGTGTGCTTATTTATCTTACC 1428
 1478 ATCAGCGAGCTGGAAGTTCGCTGCGATGTTGTTGATAGGTTTGGCTGTTGCTGTTTACA 1537
 1429 CTGTTTCAATATCATCGGACAGCTAGCGCTGCTGCTGCGCGGCGACACATCATGTG 1485
 1538 CTTTTCATATCATCGGACAGTACTGCTCTGTTATCGGACCGCATATCATGTG 1594

RESULT 12
 AX009614
 LOCUS 3109 bp DNA linear PAT 06-SEP-2000
 DEFINITION Sequence 5 from Patent EP0962528.
 ACCESSION AX009614
 VERSION AX009614.1 GI:9996846
 KEYWORDS Heliothis virescens (tobacco budworm)
 SOURCE
 ORGANISM
 Heliothis virescens
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Lepidoptera; Glossata; Ditrysia;
 Noctuoidea; Noctuidae; Heliothinae; Heliothis.
 1
 Adamczewski, M.D., Schulte, T.D. and Oellers, N.D.
 Nucleic acids encoding acetylcholin-receptor subunits from insects
 Patent: EP 0962528-A 5 08-DEC-1999;

FEATURES
 source
 1..3109
 /organism="Heliothis virescens"
 /mol_type="unassigned DNA"
 /db_xref="taxon:7102"
 95..1600
 /note="unnamed protein product"
 /codon_start=1
 /protein_id="CAC07501.1"
 /db_xref="GI:9996847"
 /db_xref="FEMTREMBL: CAC07501"
 /translation="MAPLMAALALIALIPVSEQPHEKRLNLIANLYNTLTPVANE
 SEPLEVRLFLTLOOIIDVDEKNQLITNWLSEWNDYNLRWNSYEGVGLDRLTPN
 KLPKPDVLMVNSADEGDTQYTNVVRSGCLVVPQI PKSTCKMDIAFPPEDQH
 CDMKFGSWTDGNQLDLVLDEAGDLSDFITNGEWILGMPGKNITITYACCPPEYV
 DVFTTMMIRRTLYIFFNLIIVLSSMALLGFTLPDPSGKLTGVTIILLSTVFL
 NVAETPLQVSDAIPLLGTYFNCIMFVASSVLTVVLYNHYHRTADTHEPOMIKSV
 FLOWLPILRMSRPGKIKTRKTIMNTRMRELEKRSKSLANLVIDDDPRHGP
 PPNSTASTNLGPGCSIFRTDFRSPVSRPTMEDVGGGLGSHHRLHILRELOPITA
 RMKKADEAEALISDWKFAAMVVDVFLVFETLTIATVAVLLSAPHIIVQ"
 ORIGIN
 Query Match 38.6%; Score 575; DB 6; Length 3109;
 Best Local Similarity 64.6%; Pred. No. 4.6e-106;
 Matches 929; Conservative 0; Mismatches 460; Indels 48; Gaps 3;
 QY 97 CACGAGAGCGGCTACTCCACCACTTATGACCACTACAGCTACTGAGAGGCGCCCTC 156
 DB 158 CACGAGAGAGACCTCTGAAACGCTTGTCTGGCAACTACACACCTCGAGCGACCGGTG 217
 QY 157 GTCAACAGAGAGCGCCCTGACGCTCTCTTCCGCGCTCAGCTCATCAGATCATCGAC 216
 DB 218 GCCAAGAGAGAGACCGCTAGAGGTGAGGTTCGGCTTGACCTTCAGCAATCATTCAC 277
 QY 217 GTGAGCAGAGAGAACCGAGCTTTTAAACAAACATCTGGCTTAAACATAGATGGAATGAT 276
 DB 278 GTGAGCAGAGAGAACTCACTTATTAACCAATATATGGCTGTCTGAGTGAATGAC 337
 QY 277 ATGACATGAGTGGAGACACTTCAGATTTCCGGGGGTCAAGATTTAAGATGCGCACCC 336
 DB 338 TACAACTGAGGTGGAACGACAGCGAGTATGGCGGGTCAAGGACCTCAGGATCACGGCC 397
 QY 337 CACAGACTATGAAACACAGAGCTCTTATGTACAAACAGCGGAGCAAGGGTTCACAGC 396
 DB 398 AACAGTGTGGAAGCGGACGCTCTTATGTATTAATAGTGTGACGAGGGTTTGACCGG 457
 QY 397 ACGTATCCAGCAACGTTGGTGGGAAACAAACGCTCGTGTCTGTAGTGGCGCGCCGCG 456
 DB 458 ACCTACAGACCAACGTTGGTGGTGGAGCGCGCGGAGTTGCTGTAGTGGCACTGGCC 517
 QY 457 ATCTTTCAAGAGCACTCAAGATCGACATCACCTGTTCCCTTCGAGACCAACGATGC 516
 DB 518 ATATTCAGAGCAATGCAAGATGACATCGCGTGGTTTCCCTTCGAGCAACACACTGT 577
 QY 517 GAGATGAAGTTTGGCAGCTGAGCTTATGATGGTTATCAGTTGGATCTACAACATCAGGAT 576
 DB 578 GATATGAAGTTTCGGTAGCTGACATATGACGCAATCAGTTGGATCTGGTCTAAAAGAT 637
 QY 577 GAAGGGGCGGAGATATAGCAGTTTGTACGAAATGGGAGTGGAGTTAATAGGATGC 636
 DB 638 GAGGACGCGCGGATCTATCGGACTTCATAAATAATGGGAGTGGTATCTAATAGGATG 697
 QY 637 CCCGGCAAGCGCAACGAGATCTACTACAACTGTTGTCGGAGGCATATCATCGACATCAG 696
 DB 698 CCAGGCAAAAGAACACAAATACATACGCTGTGCCCCGAGCCCTACGTGAGAGTCA 757
 QY 697 TTTCGGTGGTGGATCCGAGGAGAAACGCTCTACTACTTCTTCAATCTGATCGTCCCTGCG 756
 DB 758 TTCACCATCATGATAAGAGACGAACTTGTACTACTTCTTCAACCTGATCGTCCGTCG 817
 QY 757 GTGCTCATCGCTCCATGGCTCTATTGGGGTTCACCTTGGCTGCGAGCTCCGAGAAAG 816

Db 818 GTGCTGATCTCATCGATGGCACTCCTCGGCTTCACACTGCCACAGACTCCGGAGAGAA 877
QY 817 TTGTCCTTTAGTGTGACGATATTAATCTGCTGTGTGACGCTGTTCTCAACATAGTGGCGGAG 876
Db 878 CTCACACTTGGAGTCACTATTTCTATCTGCTGACGCTGTTCTCAACCTGGTAGCGGAG 937
QY 877 ACGATGCACGACGCTCGGACGCGTCCCTTGTCTCGGACCTTCACTCAACTGCATCATG 936
Db 938 ACCCTGCCACAGGCTCTCCGACGCTATCCCCCTGTAGGAGTACTTCAATGTCATCATG 997
QY 937 TTCATGTTGGCTTCTCCGCTGCTCCACCATACTGATCTCAACTACCAACCCGCGAC 996
Db 998 TTCATGTTAGCTGCTGTGTTACTGACTGTGTTGTTACTCAATTACCACTCGAACA 1057
QY 997 GCAGACACTCAGNAATGAGTGTGTTGCTGCTGCTGCTTCAATGCTGCTGCTGCTG 1056
Db 1058 GCTGATATACATGAATGCCACAGTGGATAAAATCAGTATTTCTTACAAATGTTGCCATGG 1117
QY 1057 GTGCTGGCGATGTCACGCGCCGCTCGCG---ACGACGCGCGCGCGCGCGGTACCT 1113
Db 1118 ATACTGCGAATGTCAGGCCAGGGAAGAAGATCACAGGAAGACTATAAATGATGAACACG 1177
QY 1114 CCGCGCGGACCTGGAGCTGCGGAGCGTCTCTCAAGTCTCTTCAAGGAGCGTCTGCTC 1173
Db 1178 AGGATGAGGAGCTGGAACTGAAGGAGAGTCTGTCGAAGTCTTGTGCGGAATGTTCTA 1237
QY 1174 GACATCATGACGACTTCCGCCACCCGACGAGCGGAGC-----AGCGGAATGCTGCGG 1226
Db 1238 GATATTGATGATGACTTCAGACACAGCGCCCTCCGCTCTTAACTGACTGCTCGACCGGG 1297
QY 1227 ATACTACAGGGGGTTCAGAGAGATGCGGGGTTGCGGCGCGCACAGTGTCTTCG--- 1282
Db 1298 AATTGGGACTGGTGCTCAATATTCGACGAGTTCCTGCTGCTGCTGCTGCTGCTGCTG 1357
QY 1283 -----GTGTGACTACGAGCTCTCCCTCAT 1308
Db 1358 TCCACGATGAAGACGTGGCGGGGCTGGTAGCCACCATCGGAGCTGCACCTATA 1417
QY 1309 CTGAAGGAGATTAGATCATCAGATCAGATGCGGAAGGACGACGAGATGCGGACATT 1368
Db 1418 CTGAGAGAGCTGCAGTTTCATCAGCGCCAGGATGAAGAAGCTGATGAGGAAGCGAGCTG 1477
QY 1369 TCGCGGCACTGGAAGTTGCGCGCCATGCTGTGACAGACTGTCCTTATATCTTATACC 1428
Db 1478 ATAGCGACTGGAAGTTGCTGCAUGTTGTTGATAGTTTTCCTGCTGCTGCTGCTGCT 1537
QY 1429 CTGTTCAATATATCGCCAGCTAGCGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1485
Db 1538 CTTTTCACATATCGCGACAGTAGTGTCTGTTATCGGACCGCATATCATGCTG 1594

RESULT 13
DME554210 1683 bp mRNA linear INV 01-APR-2003
LOCUS Drosophila melanogaster mRNA for nicotinic acetylcholine receptor
DEFINITION subunit Dalpha7 (nAcRalpha-18C gene).
ACCESSION AJ554210
VERSION AJ554210.1 GI:29466436
KEYWORDS nAcRalpha-18C gene; nicotinic acetylcholine receptor subunit Dalpha7.
SOURCE Drosophila melanogaster (fruit fly)
ORGANISM Drosophila melanogaster
Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 1683)
AUTHORS Millar, N.S.
TITLE Direct Submission
JOURNAL Submitted (31-MAR-2003) Millar N.S., Department of Pharmacology, University College London, Gower Street, London, WC1E 6BT, UNITED KINGDOM
FEATURES source Location/Qualifiers
1. .1683

/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
1. .1683
/gene="nAcRalpha-18C"
26. .1654
/gene="nAcRalpha-18C"
/codon_start=1
/product="nicotinic acetylcholine receptor subunit Dalpha7"
/protein_id="CAD86936.1"
/db_xref="GI:29466437"
/translation="MSFPQHSIPEANTANGRMVYGLLLIMIPACAGPHEKRLIH
ALDNNNSLPPVNSDPLQLSFLTLMQIIDVEKNQLLITNIMLKLWNNLWLEW
NSSEFGVRLDIRPHRLKPDVYNSADDFDGTATVAVVNNNGSCILYVPPGIFK
STKIDITWPFDDQCEKMGFTGDFQDLQDEAGGDISSFITNGEWLLGVF
GKNEIYVNCPEPYIDIITAILIRKTLVFFNLIVPCVLIASMLLGLTLLPDSGE
KLSGLVTLLSLTVFLNVAETWPAISDAVPLLGKYENCIMEMVASSVSTLLVNYH
HRNPDHEMSEWIRVIFLWPLCILMORQGVGYECPPPSSSSSSASGKKQIQYH
VELKRSKSLANLVIDDDDFRNHRCSATILPHQTYRTMYRQDDGSGVPGFA
GPVVDGLRHEAISHITSSAEYELALILKELRWITQLKKEDETSITDRWKPAAV
VDRLLIIFLTIIATLAVLFSAPHFIVSGVRG"
26. .131
/gene="nAcRalpha-18C"
132. .1651
/gene="nAcRalpha-18C"
/product="nicotinic acetylcholine receptor subunit Dalpha7"
ORIGIN
Query Match 38.2%; Score 568.4; DB 3; Length 1683;
Best Local Similarity 65.0%; Pred. No. 9.9e-105;
Matches 979; Conservative 0; Mismatches 416; Indels 111; Gaps 5;
QY 91 GGGTACCACAGAGCGGCTACTGCACCACTATTGACCACTACACGACTACGAGG 150
Db 131 GGACCCATGAGAGCGGCTACTCCAGCCCTTCTGGACAACTACACACGCTGGAGCGT 190
QY 151 CCCGTGCTCAACGAGACGACCCGCTGACGCTCTCCCTTCGGCTCAGCTCATGCGATC 210
Db 191 CCGTGTGTCATGAATCCGATCCATGCACTGAGCTTCGAGCTAACACCTCATGCGATT 250
QY 211 ATCGAGTGCACGAGAGAACGAGCTTTTAATAACAACATCTGGCTAAACTAGAGTGG 270
Db 251 ATCGATGTGACGAAAAGAAATCAACTGCTTATAACGAATATTTGGCTCAAAATGG 310
QY 271 AATGATATGAATTTGAGTGGAAACACTTCAGATTTTCGGCGGGTCAAAGATTTAAGAGT 330
Db 311 AACGATATGAATCTTCATGGAAATTCGAGTGAATTCGGTGTGTGGGATCTCGGATT 370
QY 331 CCACCCACAGACTATGGAAACGAGCTCTTATGTAACAACGCGCGGACGAGGTTTC 390
Db 371 CCGCCACATCGCTATGGAAACCGGATGTAATGATGTAACAAGTCCGACGAGGCTTC 430
QY 391 GACAGACGCTATCCACGAGCTGTTGGTGGGAAACGAGCTGCTGTCTGTAGTCCG 450
Db 431 GATGGAACTGACCGCAAAATGTGGTGGTTCGCAATTAATGGGAGCTGTCTGTACGTACC 490
QY 451 CCCGGATCTTCAAGAGCACTCCAGATCGACATCACCTGTTCCCTTCGACGACCAA 510
Db 491 CCAAGTATATTTAAGTCAAGTGAAGTGAATGATGCAATTCAGTGGTTTCCATTCGACATCAG 550
QY 511 CGATGCGAGATGAAGTTTGGACGCTGGACTTATGATGGTTATCAGTTGATCTTACAACTA 570
Db 551 AGATGTGAAATGAATTTGGTTCGTGGACCTACGATGGGTTTCAGTTGGACCTGCAGTTG 610
QY 571 CAGATGAAGGGGCGGAGATATAACGAGTTTTCACGATCGGAAATCGGAGTTAATA 630
Db 611 CAGGACGAAGCTGCTGGCGCAATTTCTAGCTTTATAACCAATGGCAATGGGACTTTGTTA 670
QY 631 GGAGTCCCCGGCAAGCGCAACGAGATCTACTACAACTGTTGTCGGGAGCCATACATCGAC 690
Db 671 GGTGTGCGGTTAAAGAAATGAATCTACTATATTTGTCGCCAGAACCTTATATTGAC 730


```
QY 1191 CCGCCACCCGACAGCGCAGCGCAATGCTGCCGATCTACTACAGGGGGGTGAGGAGAA 1250
Db |||||
QY 1647 TCGGCACACAATATCTGGCTCCCAAAACGGCCATTGGCTCGTCGCCAGCTTCGGTCGGCC 1706
Db |||||
QY 1251 TGGCGGGGGTTGGCGGCGCACAGTTGCTTTCGGTGTGCACTAC---GAGCTCTCCCTCAT 1307
Db |||||
QY 1707 CACAACGGTGGAGAGCATCACACGGCCATCGGCTGCAATCAAAAGATCTTCATCTAAT 1766
Db |||||
QY 1308 TCTGAAGGAGATTAGATCATCAGATCAGATGCGCAAGGACGACGAGATCGGACAT 1367
Db |||||
QY 1767 TCTTAAAGAAATTCGAATTTATTACGGCGCGGATCGCAAAAGCTGACGACGAGCGAATT 1826
Db |||||
QY 1368 TTCGCGGACTGGAGTTCCGCCCATGGTCTGTCGACAGACTGTGCCTTATTCTTTAC 1427
Db |||||
QY 1827 GATCGGCGATTGGAAGTTCGGGCAATGGTGTGGATAGATTTTGTTAATTGTTTTCAC 1886
Db |||||
QY 1428 CCTGTTTCAATCATGCCACGGTAGCCGTGCTGCTGTCGCCGCCACACATCATGGTG 1485
Db |||||
QY 1887 GCTCTTCAAGATTATTGCAACGGTTACGGTGTGCTCTCCGCTCCGACATATCGTG 1944
Db |||||
```

Search completed: May 8, 2004, 02:41:22
Job time : 5903.18 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 10:08:00 ; Search time 598.456 Seconds
(without alignments)
10562.710 Million cell updates/sec

Title: US-09-303-232-3_COPY_335_1822

Perfect score: 1488
Sequence: 1 atggggggggggcgccgcg.....cgccacacatcatgtgtcg 1488

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseq 29Jan04:*

1: geneseqn1980s:.*
2: geneseqn1990s:.*
3: geneseqn2000s:.*
4: geneseqn2001as:.*
5: geneseqn2001bs:.*
6: geneseqn2002s:.*
7: geneseqn2003as:.*
8: geneseqn2003bs:.*
9: geneseqn2003cs:.*
10: geneseqn2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	ID	Description
1	1488	100.0	3700	3	Aaz24476 H. viresc
2	575	38.6	3109	3	Aaz24477 H. viresc
3	512.8	34.5	2886	3	Aaz24475 D. melano
4	470	31.6	936	4	Ab113733 Drosophil
5	458.2	30.8	1540	4	Ab107231 Drosophil
6	405.6	27.3	1509	4	Aac90380 Wild-type
7	405.6	27.3	1876	2	Aat48239 Neuronal
8	405.6	27.3	1876	8	Adal0864 Human neu
9	405.2	27.2	1876	2	Aav12197 Human neu
10	405.2	27.2	1876	6	ABs54875 Human neu
11	405.2	27.2	1876	6	ABv73248 Human neu
12	404	27.2	1509	3	Aac58395 Human PRO
13	404	27.2	1509	4	Aac90385 Mutant hu
14	404	27.2	1509	4	Aac90386 Mutant hu
15	402.4	27.0	1509	4	Aac90387 Mutant hu
16	394	26.5	1590	2	Aav44687 V274T var
17	386.8	26.0	1964	6	Abz11298 Human pol
18	327.4	22.0	2769	2	Aat59196 Neuronal
19	312.2	21.0	2101	2	Aat59197 Neuronal
20	309.2	20.8	2277	2	Aav12199 Human neu
21	309.2	20.8	2277	2	Aaq90387 Alpha 2 s
22	309.2	20.8	2277	6	ABs54870 Human neu
23	309.2	20.8	2277	6	ABv73243 Human neu

24	309.2	20.8	2664	2	AAT48235	Rat48235 Neuronal
25	309.2	20.8	2664	6	ABK92165	Adk92165 Prostate
26	309.2	20.8	2664	8	ADAl0854	Adal0854 Human neu
27	309.2	20.8	2664	9	ADC71170	Adc71170 Human 205
28	309.2	20.8	2666	9	ADC71168	Adc71168 Human 205
29	302.4	20.3	803	4	ABL07799	Ab107799 Drosophil
30	292	19.6	1908	2	AAT48236	Rat48236 Neuronal
31	292	19.6	1908	8	ADAl0856	Adal0856 Human neu
32	290	19.5	1416	4	AAC90382	Aac90382 Chimeric
33	289.8	19.5	1771	9	ADB78675	Adb78675 Human nic
34	289.8	19.5	1771	9	ADB78664	Adb78664 Human nic
35	288.2	19.4	1771	9	ADB78674	Adb78674 Human nic
36	288.2	19.4	1771	9	ADB78676	Adb78676 Human nic
37	286.6	19.3	1771	6	ADB26345	Ad26345 Human mut
38	286.6	19.3	1771	9	ADB78665	Adb78665 Human nic
39	286.6	19.3	1771	9	ADB78666	Adb78666 Human nic
40	286.6	19.3	1771	9	ADB78678	Adb78678 Human nic
41	286.2	19.2	1809	5	AAS91552	Ras91552 DNA encod
42	286.2	19.2	2082	2	AAT59527	Rat59527 Alpha4 su
43	285.2	19.2	1654	6	ABS54871	Ab54871 Human neu
44	285.2	19.2	1654	6	ABV73244	Abv73244 Human neu
45	285.2	19.2	1757	2	AAV12200	Aav12200 Human neu

ALIGNMENTS

RESULT 1

AAZ24476

ID AAZ24476 standard; cDNA to mRNA; 3700 BP.

XX

AC AAZ24476;

DT 17-FEB-2000 (first entry)

DE H. virescens acetyl-choline receptor DNA from clone Hva7-1.

XX

KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;
KW neurotransmission; plant protection agent; conductance; AChR; ds.
XX Heliothis virescens.

XX

PH Key

CDS Location/Qualifiers

FT 335..1825

FT /*tag= a

FT /product= "acetyl-choline receptor"

XX

PN DE19819829-A1.

XX

PD 11-NOV-1999.

XX

PF 04-MAY-1998; 98DE-01019829.

XX

PR 04-MAY-1998; 98DE-01019829.

XX

PA (FARB) BAYER AG.

XX

PI Adamczewski M, Oellers N, Schulte T;

XX

WPI; 2000-014207/02.

DR P-PSDB; AAY50815.

XX

FT New nucleic acid encoding a nicotinic acetylcholine receptor from

insects, used to identify potential insecticides.

XX

PS Claim 1a; Page 14-17; 26pp; German.

XX

CC This invention describes a novel nucleic acid (NA) encoding a nicotinic

acetyl-choline receptor (I) from insects which can be used as an

insecticide. Inhibitors of (I) interfere with neurotransmission. (I)

CC (also vectors containing it, its regulatory regions, and antibodies

directed against (I)-encoded proteins) are used to screen for: (a) plant

protection agents that alter conductance of AChR, potentially useful as

insecticides, or (b) genes which encode polypeptides that are involved in formation of functionally related AChR in insects. (I) are also used to isolate and characterize the specified regulatory regions and for recombinant production of (II). This sequence encodes an acetyl-choline receptor isolated from *Heliothis virescens*

Query Match 100.0%; Score 1488; DB 3; Length 3700;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 1488; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 ATGGGCGGGGCGGCGCGCTCCGACCTTGGCGGCGCGCGGCGCTGCTGCTGCTG 60
Db 335 ATGGGCGGGGCGGCGCGCTCCGACCTTGGCGGCGCGCGGCGCTGCTGCTGCTG 394

Qy 61 TGCCTGCTCTGGCGAGGGGGGCGACGCTGGGGTACCAAGAGAGGGCTACTCGACAC 120
Db 395 TGCCTGCTCTGGCGAGGGGGGCGACGCTGGGGTACCAAGAGAGGGCTACTCGACAC 454

Qy 121 CTATTGACCACTAACAGTACTGGAGAGGCGCGTCTCAACGAGAGCGACCGCTCGAG 180
Db 455 CTATTGACCACTAACAGTACTGGAGAGGCGCGTCTCAACGAGAGCGACCGCTCGAG 514

Qy 181 CTCTCTTGGCGCTCAGCTCATCGACATCATCGACGTGGAGAGAGAAACAGCTTTTA 240
Db 515 CTCTCTTGGCGCTCAGCTCATCGACATCATCGACGTGGAGAGAGAAACAGCTTTTA 574

Qy 241 ATAACAACATCTGGCTAAACTAGAGTGGATGATATGAATTTGAGGTGGAACACTTCA 300
Db 575 ATAACAACATCTGGCTAAACTAGAGTGGATGATATGAATTTGAGGTGGAACACTTCA 634

Qy 301 GATTTCGGGGGGTCAAGAGTTTAAGAGTGCACCCACAGACTATGGAACAGAGCTC 360
Db 635 GATTTCGGGGGGTCAAGAGTTTAAGAGTGCACCCACAGACTATGGAACAGAGCTC 694

Qy 361 CTATTGTAACAACAGCGGGGAGAAAGGTTGACAGACAGTATCAACAGAGCTGGTGTG 420
Db 695 CTATTGTAACAACAGCGGGGAGAAAGGTTGACAGACAGTATCAACAGAGCTGGTGTG 754

Qy 421 CGGAGAACGGCTCGTCTGTAGTGGCGGCGCGCGGATCTTCAAGAGACCTGCAAGATC 480
Db 755 CGGAGAACGGCTCGTCTGTAGTGGCGGCGCGGATCTTCAAGAGACCTGCAAGATC 814

Qy 481 GACATCACTGGTTCCTCTCGACGACCAAGATGCGAGATGAAGTTTGGCAGCTGGACT 540
Db 815 GACATCACTGGTTCCTCTCGACGACCAAGATGCGAGATGAAGTTTGGCAGCTGGACT 874

Qy 541 TATGATGGTTATCAGTTGGATCTTAACTACAGGATGAAGGGCGGAGATATAAGCAGT 600
Db 875 TATGATGGTTATCAGTTGGATCTTAACTACAGGATGAAGGGCGGAGATATAAGCAGT 934

Qy 601 TTTGTACGAATGCGAATGGAGTTAATAGAGTCCCGGCAAGCGACGAGATCTAC 660
Db 935 TTTGTACGAATGCGAATGGAGTTAATAGAGTCCCGGCAAGCGACGAGATCTAC 994

Qy 661 TACAACCTGTTGTCGGAGGCATACATCGACATCACGTTTGGGTGGTGTATCCGGAGAAA 720
Db 995 TACAACCTGTTGTCGGAGGCATACATCGACATCACGTTTGGGTGGTGTATCCGGAGAAA 1054

Qy 721 ACGCTCTACTACTTCTTCAATCTGATCGTGGCTGCGTGTCTATCGGCTCCANGCTCTA 780
Db 1055 ACGCTCTACTACTTCTTCAATCTGATCGTGGCTGCGTGTCTATCGGCTCCANGCTCTA 1114

Qy 781 TTGGGGTTTCACTTGGCTCCAGACTCCGGAGAAAGTTTCTTAGTGTGACGATATTA 840
Db 1115 TTGGGGTTTCACTTGGCTCCAGACTCCGGAGAAAGTTTCTTAGTGTGACGATATTA 1174

Qy 841 CTGTCGTTGAGCGTGTCTTCAACATATGGTGGGAGACGATGCCAGCGAGCTCGGAGGCC 900
Db 1175 CTGTCGTTGAGCGTGTCTTCAACATATGGTGGGAGACGATGCCAGCGAGCTCGGAGGCC 1234

Qy 901 GTGGCCTTGTGGGACCTACTTCACTGATCATGTTCATGTTGAGTGGCTTCTCGCTCGTC 960

Db 1235 GTGCCCTTGTCTGGCACCTACTTCAACTGATCATGTTCATGGTGGCTTCTCCGTCTC 1294
Qy 961 TCACCACTACTGATCTCTCACTACCAACACCGGACGACACTACGAAATGAGTGAT 1020
Db 1295 TCACCACTACTGATCTCTCACTACCAACACCGGACGACACTACGAAATGAGTGAT 1354
Qy 1021 TGGATTCGTTGCGTTCCTTTATTGGCTGCGTGGGTGCTGCGCATGTACGCGCCCGC 1080
Db 1355 TGGATTCGTTGCGTTCCTTTATTGGCTGCGTGGGTGCTGCGCATGTACGCGCCCGC 1414
Qy 1081 TCGGCGACGACGCGCGCGCGGCGGTACTCCGCGCGCGGACCTGGAGCTCGCGAG 1140
Db 1415 TCGGCGACGACGCGCGCGCGCGGTACTCCGCGCGCGGACCTGGAGCTCGCGAG 1474
Qy 1141 CGCTCTTCCAAAGTCGCTCTTAGCGAAACGTCGTGACATGCAATGAGAGCTTCCGCGACCCG 1200
Db 1475 CGCTCTTCCAAAGTCGCTCTTAGCGAAACGTCGTGACATGCAATGAGAGCTTCCGCGACCCG 1534
Qy 1201 CAAGCGCAGCAGCGCGCAATGCTGCCGATCTACAGGGGGGTGAGGAGAAATGGCGCGGG 1260
Db 1535 CAAGCGCAGCAGCGCGCAATGCTGCCGATCTACAGGGGGGTGAGGAGAAATGGCGCGGG 1594
Qy 1261 TTGCGGCGCACAGTTGCTTTCGCTGTCGACTACGAGCTCTCCCTCATTTCTGAAGGAGATT 1320
Db 1595 TTGCGGCGCACAGTTGCTTTCGCTGTCGACTACGAGCTCTCCCTCATTTCTGAAGGAGATT 1654
Qy 1321 AGAGTCATCAGATCAGATGCGCAAGGACGACGAAAGATGCGGAGACTTTCGCGGACTGG 1380
Db 1655 AGAGTCATCAGATCAGATGCGCAAGGACGACGAAAGATGCGGAGACTTTCGCGGACTGG 1714
Qy 1381 AAGTTCCGCCCATGGTCGTGGACAGACTGTGCCCTTATTCTTTACCTGTTTCAACAATC 1440
Db 1715 AAGTTCCGCCCATGGTCGTGGACAGACTGTGCCCTTATTCTTTACCTGTTTCAACAATC 1774
Qy 1441 ATGCCACGCTAGCGTGTGCTGTCCGCGCCACACATCATGTGTCTG 1488
Db 1775 ATGCCACGCTAGCGTGTGCTGTCCGCGCCACACATCATGTGTCTG 1822

RESULT 2

AAZ24477

ID AAZ24477 standard; cDNA to mRNA; 3109 BP.

XX AC AAZ24477;

XX DT 17-FEB-2000 (first entry)

XX DE H. virescens acetyl-choline receptor DNA from clone Hva7-2.

XX KW Acetyl-choline receptor; nicotinic; insect; insecticide; screening;

XX KW neurotransmission; plant protection agent; conductance; AChR; ds.

XX OS Heliothis virescens.

XX PN DE19819829-Al.

XX PD 11-NOV-1999.

XX PF 04-MAY-1998; 98DE-01019829.

XX PR 04-MAY-1998; 98DE-01019829.

XX PA (FARB) BAYER AG.

XX PI Adamczewski M, Oellers N, Schulte T;

XX DR WPI; 2000-014207/02.

XX DR P-PSDB; AAY50816.

XX PT New nucleic acid encoding a nicotinic acetylcholine receptor from

XX insects, used to identify potential insecticides.

xx Adamczewski M, Oellers N, Schulte T;
xx WPI; 2000-014207/02.
xx P-PSDB; AAY50814.
xx New nucleic acid encoding a nicotinic acetylcholine receptor from
xx insects, used to identify potential insecticides.
xx Claim 1a; Page 8-12; 26pp; German.
xx This invention describes a novel nucleic acid (NA) encoding a nicotinic
xx acetyl-choline receptor (I) from insects which can be used as an
xx insecticide. Inhibitors of (I) interfere with neurotransmission. (I)
xx (also vectors containing it, its regulatory regions, and antibodies
xx directed against (I)-encoded proteins) are used to screen for: (a) plant
xx protection agents that alter conductance of AChR, potentially useful as
xx insecticides, or (b) genes which encode polypeptides that are involved in
xx formation of functionally related AChR in insects. (I) are also used to
xx isolate and characterize the specified regulatory regions and for
xx recombinant production of (II). This sequence encodes an acetyl-choline
xx receptor isolated from *Drosophila melanogaster*
xx
xx Sequence 2886 BP; 942 A; 687 C; 640 G; 617 T; 0 U; 0 Other;
xx
xx Query Match 34.5%; Score 512.8; DB 3; Length 2886;
xx Best Local Similarity 62.5%; Pred. No. 5.4e-115;
xx Matches 878; Conservative 0; Mismatches 502; Indels 24; Gaps 4;
xx
xx 91 GGGTACACGAGAGCGCTACTCCACCACTATTGGACCACTACAACGTAAGAGG 150
xx 1296 GGATATCATGAAAGAGACTGTACAGATCTTTTGGATCCTTATAATACACTAGAAGT 1355
xx 151 CCCGTCCTACGAGAGCGCCCTGCGAGTCTCTTCGCGCTCAAGCTACGATC 210
xx 1356 CCGGTCCTCAATGAATCGGACCCGTTTAAATTAAGCTTTGGTTTAACTTTAATGCAATT 1415
xx 211 ATCGACGTGACGAGAGAACCCAGCTTTTAAATCAACAACATCTGGCTAAACTAGAGTGG 270
xx 1416 ATCGATGTGACGAGAAATCAATTTGCTAGTCACTAATGTGTGTTAAACTGGAGTGG 1475
xx 271 AATGATATGAATTTAGAGTGAAACATTTAGATTTCCGCGGGGTCAAGATTTAAGAGTG 330
xx 1476 AACACATGAATCTCCGCTGGAAACACCTCCGACTATGCGGAGTTAAGGATCTCGAATA 1535
xx 331 CCACCCACAGACTATGAAACACAGAGCTCTTATGTACACAGCGCGGAGAGGCTC 390
xx 1536 CCGCGCATCGATCTGGAACCGGAGCGTGTGATGTACACAGTGGGATGAGGGAATTT 1595
xx 391 GACAGCAGTATCCAAAGAGCGTGGTGTGGGAAACACCGCTCGTGTGTAGTGGCG 450
xx 1596 GACGGCACCTACGACAGCGAGTGGTGTGGGAAACACCGCTCGTGTGTATAGTTCG 1655
xx 451 CCGGGATCTTCAAGAGACCTGCAAGATGACATCACTGTTCCCGTTCGACGACAA 510
xx 1656 CCGGGATCTTCAAGTCCAGTCAAGATCGACATCGACATCGTGGTTCGCTTCGATGAC 1715
xx 511 CGATGCGAGATGAATTTGCGAGCTGACATTTATGATGATGATGATGATGATGATGAT 570
xx 1716 CCGTGCAGATGAATTTGCGAGTGGACCTACACGAGTTCCAGCTGGATTTACAATTA 1775
xx 571 CAGATGAAGGGGCGGAGATATAAGCAGATTTTGTCAAGATGCGAATGGAGTTAATA 630
xx 1776 CAAGATGAATTTGCGGTGATATCAGCAGTTACGTCTCAACGCGAGTGGAACTACTG 1835
xx 631 GGAGTCCCGCAGACCGCAGAGATCTACTAATCTGTTGTCGGAGGCCATACATCGAC 690
xx 1836 GGTGTGCCCGCAACCGTAACGAGATCTATTACAACTGCTGCCCGAACCCCTATATAGAC 1895
xx 691 ATCAGTTTGGGTGGTGTATCCGAGGAAACCGCTTACTACTTCTCAATCTGATCGTG 750
xx 1896 ATCACTTTCGCATCATCCGCGGACGACACTGTACTTATTTCTTCAACCTGATATA 1955

Qy 751 CCCTGCGTGTCTATCGCTCCATGGCTCTATTGGGTTTCACTTGGCTTCCAGATCCGGA 810
Db 1956 CTTTGTGTACTGATTTGCTCCATGGCTTCTCGATTCAGCTTCCCTGCGGAGATTCGGGT 2015
Qy 811 GAAAAGTTGTCTTTAGGTGTGACGATATTACTGTCTGTGACGGTGTCTCTCAACATGGT 870
Db 2016 GAAAATTTATCGCTGGGTGTATCCATCTTGTCTGTGACCGGTGTCTCTGAATATGTT 2075
Qy 871 GCGGAGACGATCCAGCGACGTGGAGCGCGTGGCTTGTCTGCGACCTACTTCAATGTC 930
Db 2076 GCGGAGACATGCGCGTACTTCCGATGCGGTGCCATTTG-----TGGATACGC 2123
Qy 931 ATCATGTTTCATGTTGGCTTCTCCGTCGTCTCCACCATACTGATCTCACTACACACAC 990
Db 2124 ATCGTGTTTTG--TGTGGCTGCATGGATATTGCGAATGAGTGGCCAGGACGCGC 2181
Qy 991 CGGCACGACACACTCAAGAAATGAGTGTGATTTGGATTCGTTCGCTGTCTTTTATGGCTG 1050
Db 2182 TGATCTCTAGAGTTCCCGACCAACGCTTGTTCGGACACAT-----CCTCCGAGCGGAAGCAC 2237
Qy 1051 CCGTGGGTGTCTGCGCATGTACGCGCGCTCGCGAGACGCGCGCGCGCGCGGCGGTA 1110
Db 2238 CAGATACTCTCCGACGTTGAGCTGAAAGAGCGCTCGTCCGAATCGCTCTGGCCCAACGTA 2297
Qy 1111 CCTCCGCGCGGACCTGGAGCTGCGGAGCGCTCTCCAAAGTCGCTCCTAGCGAACCTG 1170
Db 2298 CTAGACATCGATGATGACTTCCGGCACAATTTGTCGCCCATGACCGCGCGGAAACACTG 2357
Qy 1171 CTGACATCGATGACGACTTCCGCGCACCCGCAAGCGAGCGACGCGCAATGCTCCGATAC 1230
Db 2358 CCACACAACCGCGCTTCTATCGCACGGTTTATGGAACAAGCGGACGATGGCAGCATTTGG 2417
Qy 1231 TACAGGGGGGTGAGGAGAAATGGCGGGGTTCGCGGCGCACAGTTCCTTC-----GGT 1284
Db 2418 CCAATTGCGAGACCCCGAATGCGGATGCGGTTCACCCATCATAGTGCATCAAAATCATCA 2477
Qy 1285 GTGACTATCAGAGCTCTCCCTCAITCTGAAGAGATTAGAGTCAATCAGATCAGATGCGC 1344
Db 2478 ACTGAATATGAATTAGGTTTAACTTTAAAGGAAATTCGCTTTTAACTGATCAGTACGT 2537
Qy 1345 AAGGACGACGAAGATGCGGACATTTCCGCGGACCTGGAAGTTTCGCGGCATGCTCGTGAC 1404
Db 2538 AAAGATCAGAGTGCATGACATTTGCCAATGATTTGGAATTTGAGCTATGCTCGTTGAC 2597
Qy 1405 AGACTGTCCTTATTATCTTTTACCTGTTCCAAATCATCGCCAGCTAGCGCTGCTGCTG 1464
Db 2598 AGACTGTCCTTATTATCATATTACAAATGTTTCCAAATATTAGCCAAATAGCTGTACTA 2657
Qy 1465 TCCGCGCCACACATCATGTTGTCG 1488
Db 2658 TCGGACCAACATATTATTTGTCG 2681

RESULT 4
ABL13733
ID ABL13733 standard; cDNA; 936 BP.
XX ABL13733;
AC ABL13733;
DT 26-MAR-2002 (first entry)
XX
XX *Drosophila melanogaster* expressed polynucleotide SEQ ID NO 35681.
DE *Drosophila*; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX *Drosophila melanogaster*.
XX WO200171042-A2.
XX
XX 27-SEP-2001.
XX
XX 23-MAR-2001; 2001WO-US009231.

[illegible]

Db	1022	GGTTCTGCGAATGAAGAGGCCCGGAGGACAAAGGTGCGCCCGGCTGCCAGCACAGC	1081
Qy	1115	CGCGC-----CCGGAAGCTGAGCTGCGCGAGCGCTCTCTCAAGTTCGCTCTAGCGA	1165
Db	1082	AGCGCGCTGCAGCTGCGCCAGTGTGAGATGAGCGCGTGGCGCGCGCCGCGCAGCA	1141
Qy	1166	ACGTGCTGCACATCATGATGACATTCCTCGGCCACCCGCGAAGCGCAGCAGCCGGAATGCTGCC	1225
Db	1142	ACGGGAACCTGCTGTATCATCGGCTTCGCGCGCCTTGAGACGGCGTGCATGTGTCCCGACCC	1201
Qy	1226	GATACTACAGGGGGGTGAGAGAAATGCGCGCGGGTTGGCGGGCGCACAGTTGCTTCGGTG	1285
Db	1202	CCGACTCTGGGGTAGTGTGTGGCGCATGGCCTGCTCCCCACGCACATGAGCACCTCC	1261
Qy	1286	TCGACTACGAGCTCTCCCTC-----ATTCTGAAGGAGATTA	1321
Db	1262	TGCAAGCGGGCAACCCCGGAGGGGACCCGGACTTGGCCAGATCTGGAGGAGTCC	1321
Qy	1322	GAGTCATCAGATCAGATGGCAGGACAGCAGAGATGGGCATTTCCGCGGACTTGGGA	1381
Db	1322	GCTACATTGCCAATTCGCTTCGCTGCCAGACGAAAGCGAGCGGCGTGTGCAGCGAGTGA	1381
Qy	1382	AGTTCCGCGGCATGTCGTGTGCACAGACTGTCCTTATCTTTACCTGTTCAATCA	1441
Db	1382	AGTTCCGCGGCTGTGTGTGGACCGCCTGTGCTCATGGCCTTCTCGGCTTCCACCATCA	1441
Qy	1442	TCGCGACGCTAGCCGCTGTGTCTCCGCGCACACATCATGGTGTC	1487
Db	1442	TCTGCACATCGGCATCTGATGTGCGCTCCAACTTCGTGGAGGC	1487
RESULT 7			
AAT48239			
ID	AAT48239	standard; DNA; 1876 BP.	
XX			
AC	AAT48239;		
DT	09-APR-1997	(first entry)	
XX			
DE	Neuronal nicotinic acetylcholine receptor alpha-7 subunit DNA.		
XX			
KW	Neuronal nicotinic acetylcholine receptor; nAChR; neurotransmitter;		
XX	ligand-gated receptor; ds.		
OS	Homo sapiens.		
XX			
FT	Key	Location/Qualifiers	
FT	CDS	73..1581	
FT		/*tag= a	
XX			
PN	W09641876-A1.		
XX			
PD	27-DEC-1996.		
XX			
PF	07-JUN-1996;	96WO-US009775.	
XX			
PR	07-JUN-1995;	95US-00484722.	
XX			
PA	(SIBI-) SIBIA NEUROSCIENCES INC.		
XX			
PI	Elliot KJ, Harpold MM;		
XX			
DR	WPI; 1997-065463/06.		
XX			
DR	P-PSDB; AAW09025.		
XX			
PT	Nucleic acids encoding nicotinic acetylcholine receptor sub-units - used		
PT	in screening to determine the effect of drugs on the receptor.		
XX			
PS	Disclosure; Page 71-73; 108pp; English.		
XX			
CC	A DNA sequence (AAT48239) codes for the alpha-7 subunit (AAW09025) of the		
CC	human neuronal nicotinic acetylcholine receptor (nAChR). Host cells, esp.		

Nucleic acids encoding nicotinic acetylcholine receptor sub-units - used in screening to determine the effect of drugs on the receptor.

Disclosure: Page 71-73; 108pp; English.

A DNA sequence (AA748239) codes for the alpha-7 subunit (AAW09025) of the human neuronal nicotinic acetylcholine receptor (nAChR). Host cells, esp.

QY	35	GC	CCGCGGGCCTGCTGCTGCTGTGTCCTGTCTGTGCCGAGGGGGGCA	CGCTGCGGGT	94
DB	83	CG	CGGAGGCGTGTGCTGCGTGTGCGCGCGTGTGCTGCTGCACTGTCCCTGCAAGCG	142	
QY	95	ACA	CAGAGAAGCGGCTACTGCACCACCTATTGGACCACTACAACGTACTTGAGAGGCGCG	154	
DB	143	AGT	TCCAGAGAGCTTTACAAGAGCTGTGTCAAGAACTCAATCCCTTTGGAGAGGCGCG	202	
QY	155	TG	GTCAACGAGAGCGACCCGCTGCAGCTCTCTTTGGCCTCACGCTCATGCAGATCATCG	214	
DB	203	TGG	CCAATGATCGCAACCACTACCCGCTACTTCTCCCTGAGCCTCTCTGCAGATCATGG	262	
QY	215	AC	TGGACGAGAGAACAACAGCTTTTAACAACAACATCTGGCTAAACATAGAGTGGATG	274	
DB	263	AC	TGGATGAGAGAACAACAAGTTTAAACCAACAATTGGCTGCAAAATGTCTTTGGACAG	322	
QY	275	ATA	TGAACCTTGAGGTGGAACACTTTCAGATTTCGGCGGGGTCAAAAGATTTAAGAGTGCAC	334	
DB	323	AT	CATATTACAGTGGAACTGTGTCAAGAAATCCAGGGGTCAAGACTGTTCGTTTCCAG	382	
QY	335	CCA	CAGACTATGGAACCAAGACGTCCTTAATGTAAACAACGCGGAGCAAGGGTTCGAACA	394	
DB	383	ATG	CCAGATTTGGAACCAAGACATCTTCTCTATAACAGTGTGTATGAGCGCTTTGACG	442	
QY	395	GC	AGTATCCAAAGAAAGTGGTGGCGGAACAACGCTCGTGTCTGCTACGTGCCGCGCG	454	
DB	443	CCA	CATTCCACCTAAAGTGTGGTGAATTCCTTCGGGCAATGCGACGACTCGCTCCAG	502	
QY	455	GC	ATCTTCAAGAGCACCTGCAAGATGCACATCACCTGGTTCCCTTCGAGGACCAACGAT	514	
DB	503	GC	ATTTCAAGAGTTCTCTGTACATCGATGTAGCTGTTCCTTTGATGTGCAGCACT	562	
QY	515	GCG	AGATGAAGTTTGGCAGCTGGACTTATGATGGTTATCAGTTGGATCTCAACTACAGG	574	
DB	563	GCA	AACTGAAGTTTGGTCTCGTCTTACGAGGCTGGTCTTGATCTGCAGATGCAGG	622	
QY	575	AT	GAGGGGCGGAGATATAAGCAGTTTGTTCAGAAATGCGAATGGAGTGAATAGGAG	634	
DB	623	A-----	GGCAGATATCAGTGGCTATATCCCAATGAGAAATGGGACCTAGTGGAA	673	
QY	635	TCC	CGGCAAGCGCAACGAGATCTATCAACTGTGTTCGCGAGCCATAATCGACATCA	694	
DB	674	TCC	CGGCAAGAGAGGTGAAGGTTCTATGAGTGTCTCAAAAGAGCCCTACCCGATGTCA	733	
QY	695	CG	TTTGGGTGTGATCCGAGGAGAAACGCTCTACTACTTCTTCAATCTGATCGTGCCT	754	
DB	734	CC	TTCAAGTGAACATGCGCGCAGGACGCTTACTATGCGCTCAACCTGCTGATCCCT	793	
QY	755	GC	GTGCTCATCGCCTCCATGGCTCTATTGGGGTTCACCTTGGCTCCAGACTCCGGAGAAA	814	
DB	794	GT	GTGCTCATCTCGGCCCTCGCTGCTGGTGTCTCTGCTTCTGTGCAATTCGGGGAGA	853	
QY	815	AG	TGCTTTAGGTGTGACGATTAATCTGCTGTGACGGTGTTCCTCAACATGTTGGCGG	874	
DB	854	AGA	TTTCCCTTGGGGATAACAGTCTTACTCTCTTTACCGTCTTCACTGCTCTGTGGCTG	913	
QY	875	AGA	CGATGCCAGACAGCTCGGAGCGCGCTTGTCTGGGCACTACTTCAACTGATCA	934	
DB	914	AGA	TATGCGCGCAACATCCGATCGGTACCAATTGATAGCCAGTACTTCGCCAGACCA	973	
QY	935	TG	TTATGTGTGGCTTCTCGTGTCTCCACCAATCTATGATCTCTCAACTACCAACACCGCG	994	

Db 1274 CCGACTCTGGGTTAGTGTGTGGCGCATGGCCCTGCTCCCCACGACGATGAGCACCTCC 1333
QY 1286 TCGACTACGAGCTCTCCCTC-----ATCTGAAGGAGATTA 1321
Db 1334 TGCAGGGGGCAACCCCGAGGGGACCCGGACTTGGCCAAAGATCTGGAGGAGGTCC 1393
QY 1322 GAGTCATCACAGATCAGATGGCGAAGGACGACGAGATGCGGACATTTTCGGCGCACTGGA 1381
Db 1394 GCTACATTGCCAATCGCTTCGCTCCAGGACGAAAGGAGCGGTCTGCAGCCAGTGA 1453
QY 1382 AGTTTCGGCGCATGGTGTGACAGACTGTGCTTATTAUCTTACCCCTGTTCAATCA 1441
Db 1454 AGTTTCGGCGCTGTGTGTGACGCGCTGTGCTCATGGCCTTCTCGGTCTTTCACCATCA 1513
QY 1442 TCGCCAGCTAGCGGTGCTGTGTCGCGCCGACACATCATGTTGTC 1487
Db 1514 TCTGCACATCGGCATCCTGATGCTCGGTCCCAACTTCGTGGAGGC 1559

RESULT 11
ABV73248
ID ABV73248 standard; cDNA; 1876 BP.
XX AC ABV73248;
XX 22-JAN-2003 (first entry)
XX Human neuronal NACHR alpha7 subunit encoding cDNA.
XX Human; neuronal; nicotinic acetylcholine receptor; NACHR; drug screening;
KW immunochemistry; NACHR alpha7 subunit; gene; ss.
XX Homo sapiens.

Key Location/Qualifiers
CDS 73..1581
FT /*tag= a
FT /product= "NACHR alpha7 subunit"
FT /note= "neuronal nicotinic acetylcholine receptor"
XX W0200259266-A2.
XX 01-AUG-2002.
XX 29-OCT-2001; 2001WO-US050985.
XX 01-NOV-2000; 2000US-00703951.
XX (MERI) MERCK & CO INC.
XX Gillespie A, Claeps BO, Chavez-Noriega LE, Siegel R, Elliott KJ;
XX WPI; 2002-698532/75.
XX P-PSDB; ABB82435.
XX Cell comprising nucleic acids encoding human alpha and beta subunits of
PT neuronal nicotinic acetylcholine receptors, useful for in vitro screening
PT of a drug substance in a test system specific for humans.
XX Example; Page 129-130; 143pp; English.

XX The invention relates to a suitable host cell transfected with an
CC isolated nucleic acid molecule comprising a sequence of nucleotides or
CC ribonucleotides that encodes at least one alpha or beta subunit of a
CC human neuronal nicotinic acetylcholine receptor (NACHR). The compositions
CC and methods of the present invention, which provide a means to prepare
CC synthetic or recombinant receptors and receptor subunits that are
CC substantially free of contamination from many other receptor proteins,
CC are useful for observing the effect of a drug substance on a particular
CC subtype to perform in vitro screening of the drug substance in a test
CC system that is specific for humans. The antibodies can be used in
CC immunochemistry and for diagnostic and therapeutic applications. The
CC present sequence represents a human neuronal NACHR alpha7 subunit

CC encoding cDNA
XX Sequence 1876 BP; 369 A; 553 C; 530 G; 423 T; 0 U; 1 Other;
SQ Query Match 27.2%; Score 405.2; DB 6; Length 1876;
Best Local Similarity 56.5%; Pred. No. 9.1e-89;
Matches 839; Conservative 1; Mismatches 604; Indels 42; Gaps 3;
QY 35 CGCCCGCGGCGCTCTGCTGCTGTGCTGCTCTGCGGCGAGGGGCGACGCTGCGGCT 94
Db 83 CGCCGGAGGCGCTGTGCTGGCGCTGGCGCGCTCGCTCTGCACTGCTCCCTGCAAGCG 142
QY 95 ACCACGAGAGCGGCTACTGACCACTATTGGACCACTACAACTACTGGAGAGGCGCG 154
Db 143 AGTTCCAGAGGAGCTTTTACAAGAGCTGGTCAAGAACTACAATCCCTTGGAGAGCGCG 202
QY 155 TCGTCAACGAGAGCGACCCGCTGCAGCTCTCTTGGGCGCTCACGCTCATGAGATCATCG 214
Db 203 TGGCCAATGATCGCAACCACTACCGCTCTACTTCTCCCTGAGCGCTCTGCAATCATGG 262
QY 215 ACGTGGACGAGAGCAACGCTTTTATAACAACAATCTGCTAAACTAGAGTGGATG 274
Db 263 ACGTGGATGAGAGAACCAAGTTTAAACCAACATTTGGCTGCAAAATGTCTTGGACAG 322
QY 275 ATATGAACCTTGAGGTGGAACACATTCAGATTTTCGGCGGGGTCAAAAGATTTAAGAGTGCAC 334
Db 323 ATCACTATTACAGTGGATGTGTGAGAAATCCAGGGGTGAAGACTGTCTGTTTCCAG 382
QY 335 CCCACAGACTATGGAACCGAGCGTCTTATGTAACAACAGCGCGAGGAGGTTTCGACA 394
Db 383 ATGCCAGATTTTGGAAACCGAGACTTCTTCTATAACAGTGTGATGAGCGCTTTGACG 442
QY 395 GCAGTATCCAAACGAGCGTGTGTCGGAACACGCTCGTGTCTGACGTGCGCGCG 454
Db 443 CCACATTCACACTAACGCTGTTGGTGAATTTCTTGGGCAATGCGAGTACCTGCTCCAG 502
QY 455 GCATCTTTCAAGAGCACCTGCAAGATCGACATCACTGTTTCCCTTCGACGACCAACGAT 514
Db 503 GCATTTCAAGAGTTCTCTGTACATGATGATGATGATGATGATGATGATGATGATGATGAT 562
QY 515 GCGAGATGAAGTTTGGCAGCTGAGCTTATGATGTTATCAGTTGGATCTACAACATCAGG 574
Db 563 GCAAACTGAAGTTTGGGTCCTGCTTACGAGGCTGCTCTTGGATCTGAGATGAGG 622
QY 575 ATGAAGGGGCGGAGATATAAGCAGTTTGTGACGAATGGGATGGAGTTAATAGAG 634
Db 623 A-----GGCAGATATCAGTGGCTATATCCCAATGGAGATGGGACTAGTGGAA 673
QY 635 TCCCGGCAAGCGCAACGAGATCTACTACAACCTGTTTTCGGGAGCCCATACATGACATCA 694
Db 674 TCCCGGCAAGAGGAGTGAAGGTTTCTATGAGTGTGCAAAAGAGCCCTACCCCGATGTCA 733
QY 695 CGTTTGGGTTGATCCGGAGGAAACGCTCTACTACTTCTCAATCTGATCTGCTGCGCT 754
Db 734 CTTTCAAGTGAACATGCGCGGAGGAGCTCTACTATGGCTCAACCTGCTGATCTCCCT 793
QY 755 GCGTGTCTATCGCTCCATGCTCTATTGGGTTTCACTTGGCTCCAGACTCCGAGAAA 814
Db 794 GTGTCTCATCTCGCCCTCGCCCTGCTGTGTCTCTCTGAGATTCGAGGAGA 853
QY 815 AGTTGTCTTTAGGTGTGACGATATTACTGCTGAGCGGTTTCTCAATCATGTTGCGG 874
Db 854 AGATTTTCCCTGGGGATAACAGTCTTACTCTCTTACCGCTCTTCACTGCTGCTGCGGTG 913
QY 875 AGACGATGCGAGCGCTGCGAGCGCGCTGCGCTGCTCGCACCTACTTCAACTGATCA 934
Db 914 AGATCATGCGCGCAACATCCGATTCGGTACCAATGATAGCCCACTCTCGCCAGCACA 973
QY 935 TGTTTATGTTGGCTTCTCCGCTGCTTCCACCATCTGATCTCACTCACTACACACCGCG 994
Db 974 TGATCATGTGGGCTCTCGGTGTGTGAGCGGTGATCGTGTGCTGACGTACACACACCG 1033
QY 995 ACGCAGACACTCAGAAATGAGTATGATGATGATGATGATGATGATGATGATGATGATGAT 1054

Db 131 TGGCAATGACTCGCAACCACTACCGTCTACTTCTCCCTGAGCCTCTGAGATCATGG 190
QY 215 ACGTGGACGAGAAACCAAGCTTTTAAATAACAAACATCTGGCTAAACTAGAGTGGAAATG 274
Db 191 ACGTGGATGAGAGAAACCAAGTTTTAAACCAACCAACATTTGGCTGCAAAATGCTTTGGACAG 250
QY 275 ATATGAACTTGGTGGAAACACTTCAGATTTTCGGCGGGGTCAAGATTTAAGAGTGCCAC 334
Db 251 ATCACTATTTACAGTGAATGTGTCAGAAATATCAGGGGTGAAGACTGTTGCTTTCCAG 310
QY 335 CCCACAGACTATGGAAACCAAGACGCTCTCTATGTACAACAGCGCGCAAGGGTTTCGACA 394
Db 311 ATGGCCAGATTTGGAACACAGACATCTTCTCTATACAGTGTGATGAGCGCTTTGACG 370
QY 395 GCAGGTATCCAAAGAACGTTGGTGTGCGGAACAAACCGCTCGTGTCTGTACGTGCGCGCCG 454
Db 371 CCACATCCACACTAAACGTTGGTGAATTTCTCTGGGCATTTGCCAGTACCTGCTCCAG 430
QY 455 GCATCTTCAAGAGCACCTGCAAGATCGACATCACCTGGTTCCCTTTCGAGCAACAGAT 514
Db 431 GCATATTCAGAGTTCCTGCTACATCGATGTACGCTGGTTTCCCTTTGATGTGAGCACT 490
QY 515 GCAGATGAAGTTTGGCAGCTGGACTTATGATGGTTATCAGTTGGATCTACAACACAGG 574
Db 491 GCAAACTGAAGTTTGGGTCTGGTCTTACGAGGCTGGTCTCTGGATCTGAGATGCAGG 550
QY 575 ATGAAGGGGGGAGATATGAAGCAGTTTCTACGAATGCGGAATGGAGTTAATAGGAG 634
Db 551 A-----GGCAGATATCAGTGGCTATATCCCAATGAGAATGGACTAGTGGGAA 601
QY 635 TCCCGGCGAGCGCAACGAGATCTACTACACTGTTGTCCGGAGCCATATACGACATCA 694
Db 602 TCCCGGCGAAGAGTGAAGGTTCTATGAGTGTGCAAGAGCCCTACCCCGATGTCA 661
QY 695 CGTTTGGGTGGTGTATCCGGAGGAAACCGCTCTACTACTTCTTCAATCTGATCGTGCCCT 754
Db 662 CCTTCACAGTGACATGCGCGCGCAGCGCTCTACTATGCTGCTCAACTGCTGATCCCCA 721
QY 755 GCGTGTCTATCGGCTCCAGGCTCTATTTGGGGTTACCTTTGCCTCCAGACTCCGGAGAAA 814
Db 722 GTGTGTCTCATCTCGGCTCGGCTGCTGGTGTCTCTGCTTCTGTCAGAGATTCGGGGAGA 781
QY 815 AGTTGTCTTTAGGTGTGACATATTACTGTGCTGAGGTGTTCTTCAACATGGTGGCGG 874
Db 782 AGATTTCCCTGGGATTAACAGTCTTACTCTCTTACCGTCTTCACTGCTGCTGCTGCTG 841
QY 875 AGACGATGCCAGGACGCTCGGAGCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 934
Db 842 AGATCATGCCCGCAACATCCGATTCGGTACCAATTGATAGCCAGTACTTCGCCAGCACCA 901
QY 935 TGTTCATGTGGCTTCTCGTGTCTCCACCATATGATTCCTCAACTACCAACACCGGC 994
Db 902 TGATCATCGTGGGCTCTCGGTGTGGTGAAGCGGTGATCGTGTGCTGCTGCTGCTGCTGCTG 961
QY 995 ACGCAGACACTACGAAATGAGTGTGATTTGTTGCGGTGTTCTTTATTGGCTGCCGT 1054
Db 962 ACCCGACGGGGCAAGATGCCAAGTGGACCAAGTCACTCTTCTGAACCTGGTGGCGGT 1021
QY 1055 GGGTGCTGCGCATGTCAACGCGCGGTTCGGCGACGACGCGCGCGCGCGCGGTACCTC 1114
Db 1022 GGTTCCTGCGAATGAAGAGCCCGGGGAGGACAAAGTGTGCGCCCGGCTGCGCAGCAAGC 1081
QY 1115 CGCG-----CGGACCTGGAGCTGCGAGCGCTCTCTCAAGTGTCTCTAGCGA 1165
Db 1082 AGCGGCGCTGACGCTGGCGAGTGTGAGATGAGCGCGGTGGCGCGCGCGCGCGCGAGCA 1141
QY 1166 ACGTGTTCGACATCGATGACGATTCGCCCAACCGCAAGCGCAGCAGCGCAATGTGTC 1225
Db 1142 ACGGGAACCTGTGTATACATCGGCTTCGGCGCTTGGACGCGGTGCACTGTGTCCCGACCC 1201
QY 1226 GATACTACAGGGGGGTGAGGAGAAATGGCGCGGGTTGGCGCGCACAGTTGCTTGGTG 1285

Db 1202 CCGACTCTGGGGTAGTGTGTGGCGGCATGCGCTGCTCCCCACGCAAGATGAGCACCTCC 1261
QY 1286 TCGACTACGAGCTCTCCCTC-----ATTCTGAAGGAGATTA 1321
Db 1262 TGCAGGCGGGCAACCCCCGAGGGGACCCCGACTTGGCCAAGATCTTGGAGGAGTCC 1321
QY 1322 GAGTCATCAGATCAGATCGGCAAGGACGCAAGATGCGGACATTTCCGGGAGCTGGA 1381
Db 1322 GCTACATTTGCCAATCGCTTCCGCTGCAGGACGAAAGCGAGGCGGTCTGCAGCGAGTGA 1381
QY 1382 AGTTGGCGCGCATGCTGCTGACAGACTGTGCTTATTATCTTTACCTCTTCAACAATCA 1441
Db 1382 AGTTGGCGGCTGTGTGTGAGCGGCTGTGCTCTGCGCTTCTCGGTCTTCAACAATCA 1441
QY 1442 TCSCCAGCTAGCGGTGCTGCTGTCCGCGCCACACATCATGCTGTC 1487
Db 1442 TCTGCACCATCGGCATCCTGATGTGCGCTCCCAACTTTCGTGGAGGC 1487

Search completed: May 7, 2004, 15:01:41
Job time : 609.456 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: May 7, 2004, 14:08:32 ; Search time 3998.88 Seconds

(without alignments)
11111.850 Million cell updates/sec

Title: US-09-303-232-3_COPY_335_1822

Perfect score: 1488

Sequence: 1 atggcgggcgggcgccg.....cgccacacatcatgtgtcg 1488

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:*

1: em_estba:*

2: em_esthum:*

3: em_estin:*

4: em_estmu:*

5: em_estov:*

6: em_estpl:*

7: em_estro:*

8: em_htc:*

9: gb_esti:*

10: gb_est2:*

11: gb_htc:*

12: gb_est3:*

13: gb_est4:*

14: gb_est5:*

15: em_estfun:*

16: em_estcom:*

17: em_gss_hum:*

18: em_gss_inv:*

19: em_gss_pln:*

20: em_gss_vrt:*

21: em_gss_fun:*

22: em_gss_mam:*

23: em_gss_mus:*

24: em_gss_pro:*

25: em_gss_rod:*

26: em_gss_pbg:*

27: em_gss_vrl:*

28: gb_gssi:*

29: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
c 1	476.2	32.0	885	12	BG632919
2	320.6	21.5	1201	9	AL530299
3	309.2	20.8	2296	14	CD013901
4	281.6	18.9	1864	11	AK053497

5	281.6	18.9	2916	11	AK051730	Mus muscu
6	281.6	18.9	3126	11	AK080415	Mus muscu
7	280.4	18.8	4290	11	AK029177	Mus muscu
8	278.4	18.7	1374	29	AY406232	Mus muscu
9	275.2	18.5	1374	29	AY406230	Homo sapi
10	274.2	18.4	1466	14	CD013891	90139069
11	271.4	18.2	1436	29	AY402873	Homo sapi
12	269.6	18.1	1436	29	AY402875	Mus muscu
13	269.4	18.1	607	9	AI292581	GH151518.5
14	263.4	17.7	1436	29	AY402874	Pan trogl
15	263.4	17.7	2010	11	AK080475	Mus muscu
16	261.8	17.6	1374	29	AY406231	Pan trogl
17	258	17.3	4037	11	AK049722	Mus muscu
18	258	17.3	4046	11	AK051742	Mus muscu
19	253.8	17.1	2940	11	AK034228	Mus muscu
20	253.8	17.1	3230	11	AK083157	Mus muscu
21	245.2	16.5	908	13	BUI49265	AGENCOURT
22	244.4	16.4	3483	11	AK081254	Mus muscu
23	241.4	16.2	615	14	CB149460	K-EST0205
24	239.8	16.1	1442	29	AY402876	Homo sapi
25	236.8	15.9	1034	13	BX403124	EX403124
26	232.8	15.6	922	13	BUI15857	AGENCOURT
27	226.2	15.2	411	12	B1516733	BB160023A
28	226	15.2	1442	29	AY402877	Pan trogl
29	225.2	15.1	410	12	B1516843	BB160023B
30	222.6	15.0	1454	29	AY402878	Mus muscu
31	216	14.5	902	29	AY407186	Mus muscu
32	214.8	14.4	1603	29	AY411327	Homo sapi
33	214.6	14.4	833	14	CB245337	UI-M-FY0-
34	213.8	14.4	640	14	CB244439	UI-M-FY0-
35	213.8	14.4	1101	29	CNS0001F	Drosophil
36	212.8	14.3	1007	29	CNS000HU	Drosophil
37	208.6	14.0	1781	14	CD013889	90138814
38	208	14.0	1036	13	BX437801	BX437801
39	203.8	13.7	978	29	CNS006F9	Drosophil
40	202	13.6	720	14	CA373069	647053 NC
41	199.6	13.4	2513	11	AK033068	Mus muscu
42	198	13.3	755	10	AW914206	EST345510
43	197.2	13.3	791	14	CF739825	UI-M-HD0-
44	191.6	12.9	864	9	AUI20692	AUI20692
45	189.6	12.7	658	12	BM711715	UI-E-CL1-

ALIGNMENTS

RESULT 1
BG632919/c
LOCUS
DEFINITION
BG632919 885 bp mRNA linear EST 23-APR-2001
GH16126.3prime GH Drosophila melanogaster head pOT2 Drosophila
melanogaster cDNA clone GH16126 3 similar to CG4128; Fban004128
'ion channel', located on: 2L 30D1-30E1; 04/10/2001, mRNA sequence.

ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM

BG632919

BG632919.1 GI:13758409

EST

Drosophila melanogaster (fruit fly)

Drosophila melanogaster

Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;

Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;

Ephydroidea; Drosophilidae; Drosophila.

1 (bases 1 to 885)

Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,

Lewis, S. and Rubin, G.M.

BDGP/HMMI Drosophila EST Project

Unpublished (2001)

Other ESTs: GH16126.5prime

Contact: Stapleton, N.

BDGP

Lawrence Berkeley National Lab

One Cyclotron Rd, Berkeley, CA 94720, USA

Fax: 510 486 6798

Email: <http://www.fruitfly.org/EST>, est@fruitfly.berkeley.edu

Based upon the presence of a XhoI site followed by a run of 14 or

more T residues at the beginning of the sequence, this clone was polyadenylated. The resulting Poly-T sequence has been removed. hit genomic AE003511: arm.X [18792641..19136447] estimated-cyto:18A3-18C6: 04/10/2001 Plate: GH.161 row: C column: 2 High quality sequence stop: 784.

FEATURES

Location/Qualifiers
1..885
/organism="Drosophila melanogaster"
/mol_type="mRNA"
/db_xref="taxon:7227"
/clone="GH16126"
/sex="male and female"
/dev_stage="adult"
/lab_hosts="DHS - alpha"
/clone_lib="GH Drosophila melanogaster head pot2"
/note="Organ: head; Vector: pot2; Site 1: EcoRI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pot2. Plasmid cDNA library."

ORIGIN

Query Match 32.0%; Score 476.2; DB 12; Length 885;
Best Local Similarity 73.2%; Pred. No. 8.9e-82;
Matches 610; Conservative 0; Mismatches 223; Indels 0; Gaps 0;
Qy 193 CTCACGCTCATGCAGATCATCGACGTGGACGAGAGAACACAGCTTTTAATAACAAACATC 252
Db 885 CTACACTCATGCAGATTCATGTCGACGAGAGATCACTGCTTTATACGAAATTT 826
Qy 253 TGGCTAAACTAGAGTGGATGATGATGAACCTTGAGTGGACACACTTCAGATTCGCGGG 312
Db 825 TGGCTCAATTTGGAATGGAACGATATGAATCTTCGATGGAATTCGAGTGGATTCGGTGT 766
Qy 313 GTCAAGATTTAAGAGTGCACCCACAGACTATGGAACAGACAGCTTCCTATGATCAAC 372
Db 765 GTGCGGATCTGCGAATTCGCGCACATCGCCATATGGAACCGGATGACTGATGATCAAC 706
Qy 373 AGCGCGACGAAGGTTTCGACGACACGATATCCAAACGACGTGGTGGCGGAACAAACGGC 432
Db 705 AGTCCGACGAGGCTTCGATGGACGATACGCCAACAATGTGGTTCGCAATTAATGGG 646
Qy 433 TCGGTGTGTAAGTGGCGCGCGGATCTTCAAGAGCACCTGCAAGATCGACATCACTGG 492
Db 645 AGCTGTCTGACGTACGCGCAGGATATTTAAAGTCAACGTGTAAGATCGACATTAACGTGG 586
Qy 493 TTCCCTTCGACGACCAACGATCGAGATGAAGTTTGGCAGCTGGACTTATGATGGTTAT 552
Db 585 TTTCATTCGACGATCAGATGTAAGTGAATGAAATTTGGTTCGTGGACCTAGATGGGTTT 526
Qy 553 CAGTTGGATCTACAACTACAGATGAAGGGGGCGGAGATATAAGCAGTTTTGTCAACGAAT 612
Db 525 CAGTTGACCTGCAGTTTCGAGGACGAAGCTGGTGGCGACATTTCTAGCTTTATACCAAT 466
Qy 613 GCGAATGGAGTTAATAGAGTCCCGGACGAGCGAAGAGATCTACTACAACTGTTGT 672
Db 465 GCGAATGGAGTTGTTAGTGTGCGCGGTAACGAAATCTACTATAATTTGCTGC 406
Qy 673 CCGGAGCATACATCGACATCACGTTTGGGTGTGATCCGAGGAAACGCTACTACTAC 732
Db 405 CCAGAACCTTATATGACATAACATTCGCCATTTTGAATGAGCGCAAAACGTTTACTAT 346
Qy 733 TTCTCAATCTGATCGCTCGCTGCTCATCGCTCCATCGGCTCTATTGGGGTTCAAC 792
Db 345 TTTTCAATCTGATTTGGCGTGGTACTGATCGCTCCATCGGCACTGCTAGGGTTTACA 286
Qy 793 TTGCTCCAGACTCGGAGAAAAGTTGTTTAGGTGTGAGATTAATCTGTCGTTGACG 852
Db 285 CTGCCACAGATTCGTGTGAAAAGCTTCGCTTGGAGTTACAAATCTATTATTCGCTTACA 226
Qy 853 GTGTTCTCAACATGGTGGCGAGACGATGCCAGCAGCTCGGACGCGCTGCCCTTGTCTC 912
Db 225 GTCTTCTCAACATGGTGGCGAACAATGCCGCGACCTCCGATGCGGTACCGCTGCTC 166

Qy 913 GGACCTACTTCAACTGCATCATGTTATGTTGGTCTCTCGCTGCTCCACCATACTG 972
Db 165 GGAATTTATTTCAATTCATTAATGTTTATGGTGGCTCATCAGTTGTGCAACATATT 106
Qy 973 ATCTCAACTACCAACCGCGACGACACACTCACGAAATGAGTATGGAT 1025
Db 105 GTCCTCAATATCATCATAGAAATCCAGATACGATGAATGAGTGAATGGT 53

RESULT 2

AL530299

LOCUS

DEFINITION

AL530299 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens

cDNA clone CS0DD007P05 5-PRIME, mRNA sequence.

AL530299

ACCESSION

AL530299

VERSION

AL530299.2 GI:31068132

KEYWORDS

EST.

SOURCE

Homo sapiens (human)

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 1201)

Li,W.B., Gruber,C., Jessee,J. and Polayes,D.

Full-length cDNA libraries and normalization

Unpublished (2001)

On Feb 13, 2001 this sequence version replaced gi:12793792.

Contact: Genoscope

Genoscope - Centre National de Sequencage

BP 191 91006 EVRY cedex - France

Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr

was normalized. Library was constructed by Life Technologies, a

division of Invitrogen. This sequence belongs to sequence cluster

7646.r For more information about this cluster, see

http://www.genoscope.cns.fr/

cgi-bin/cluster.cgi?seq=CS0DD007CH03QPI&cluster=7646.r. Contact :

Feng Liang Email : fliang@lifetech.com URL :

http://fulllength.invitrogen.com/ Invitrogen Corporation 1600

Faraday Avenue Genoscope sequence ID : CS0DD007CH03QPI.

Location/Qualifiers

1..1201

/organism="Homo sapiens"

/mol_type="mRNA"

/db_xref="taxon:9606"

/clone="CS0DD007P05"

/tissue_type="NEUROBLASTOMA COT 50-NORMALIZED"

/clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED"

/note="1st strand cDNA was primed with a NotI-oligo (dT)

primer. Five prime end enriched, double-strand cDNA was

digested with Not I and cloned into the Not I and EcoR V

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

sites of the pCMVSPORT 6 vector. Library was normalized."

```
Db 348 CATTGGCTCAATGCTCTGGACAGATCACTATTTACAGTGGAAATGTCAGAAATATCC 407
QY 309 CGGGGTCAAGATTTAAGAGTGCACCCACAGACTATGGAACACAGACGTCTCTATGTA 368
Db 408 AGGGGTGAAGACTGTGCTTCCAGATGCCAGATTTGGAACACAGACATCTCTCTTA 467
QY 369 CAACAGCGCGGAGAGGGTTGCAGACAGCTATCCAAACAGTGGTGGTGGGAACAA 428
Db 468 TAAACAGTGTGATGAGCGCTTTGACGCCACATTCACACATTAACGCTGTGTGTAATCTTC 527
QY 429 CGGCTCGTGTCTGTACGTGCGCGCGGCATCTTCAAGAGCACCTGCAAGATCGACATCAC 488
Db 528 TGGGCATTGCGAGTCTGCTCCAGGCATATTCAGAGTTCCTGCTACATCATGATGACG 587
QY 489 CTGGTTCCCTTCGAGACCAACGATCGAGATGAAGTTTGGCAGCTGGACTTATGATGG 548
Db 588 CTGGTTCCCTTTGATGTGAGCACTGCAAACTGAACTTTGGTCTCTGGTCTTAACGGAGG 647
QY 549 TTATCAGTTGGATCTCAACTACAGATGAAGGGGGCGGAGATATAAGCAGTTTGTTCAC 608
Db 648 CTGGTCTCTGGATCTGCAGATGAGGAGGC-----AGATATCAGTGGCTATATATCCC 698
QY 609 GAATGGCGAATGGAGTTAATAGAGTCCCGGCAAGCGCAACGAGATCTACTACAACCTG 668
Db 699 CAATGGAGATGGACCTAGTGGGATCCCGGCAAGAGAGTGAAGGTTCTATGAGTK 758
QY 669 TTGTCCGAGCCATACATCAATCAATCAGCTGTTTGGCGTGGTATCCGGAGGAAACGCTCTA 728
Db 759 CTGCAAGAGCCCTTACCCCGATGTCACCTTCAAGTGAACCATGCGCGGAGGAGCTCTA 818
QY 729 CTACTTCTTCAATCTGATCGCTCGCTGCTCATCGCTCCATGCTCATGCTCTATGGGGTT 788
Db 819 CTATGGCTCAACCTGCTGATCCCTGTGCTCATCTCCGCCCTCGCCCTGCTGGTGT 878
QY 789 CACCTTCCCTCCAGACTCCCGAG--AAAAGTTGCTTTAGTGTGACGATATTACTCTCG 846
Db 879 CTTGCTCTCTCGAGATTCGGGGGAGAGATTTCCCTKGGGATAACAGTCTTACTCTCT 938
QY 847 TTGACGGTGTCTCTCAACATGGTGGCGGAGACGATGCCAGCGTCGCGCGCGTGC 906
Db 939 CTTACCGTCTCTATGCTGCTGCTGGCTGAGATCATGCCGCAACATCCGATTGCGTACCA 998
QY 907 TTGCTCGGCACTTCTCACTGCATCATGTTCTATGTTGG 946
Db 999 TTGAAGCCCGAGTCTCTTGGCCAGCACCATGATCATCGTG 1038

RESULT 3
LOCUS CD013901
DEFINITION 90134548 single gene library Homo sapiens cdNA, mRNA sequence.
ACCESSION CD013901
VERSION CD013901.1 GI:37777431
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
REFERENCE 1. (bases 1 to 2296)
AUTHORS Jin, P., Fu, G. K., Willson, A. D., Yang, J., Chien, D., Hawkins, P. R.,
Au-Young, J. and Stuve, L. L.
TITLE PCR isolation and cloning of novel splice variant mRNAs from known
drug target genes
JOURNAL Unpublished (2003)
COMMENT Contact: Jin, P.
Incyte Corporation
3160 Porter Drive, Palo Alto, CA 94304, USA
Tel: 650 621 8639
Fax: 650 621 8965
Email: pjin@incyte.com.
FEATURES Location/Qualifiers
source 1..2296
```

```
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone_lib="Single gene library"
/notes="Vector: pDrive Cloning Vector; RT-PCR was performed using gene-specific primers flanking the open-reading frame. PCR products were subcloned into pDrive Cloning Vector and sequenced completely using M13 forward and reverse primers. Sequencing gaps were closed by re-sequencing using primers flanking the gapped areas."

ORIGIN

Query Match 20.8%; Score 309.2; DB 14; Length 2296;
Best Local Similarity 58.0%; Pred. No. 2.8e-49;
Matches 567; Conservative 0; Mismatches 408; Indels 3; Gaps 1;

QY 100 GAGAAGCGGTACTGCACCACCTATTGGACCACTACAAACGTACTGGAGAGGCCCGTCGTC 159
Db 350 GAGGACCGGCTCTTCAACACACTCTTCCGGGGCTACAAACGCTGGGGCGCCCGGTGCC 409
QY 160 AACGAGAGCAACCGCTGTCAGCTCTCTTCGGGCTCACGCTCATGTCAGATCATCGAGTG 219
Db 410 AACACTTCAGACGCTGGTGAATGTGGCTGTGGACTGTCCATCGCTCAGCTCATCGATGTG 469
QY 220 GACGAGAGAACCAAGCTTTTAATAACAACATCTGGCTAAACTAGAGTGAATGATG 279
Db 470 GATGAGAGAACCAAAATGATGACCAACAGCTGTGGCTAAACAGGAGTGGAGCATAC 529
QY 280 AACTTCAGGTGGACACTTCAGATTTTCGGGGGTCAAAAGATTTAAGAGTGCACACCCAC 339
Db 530 AAATCGCTGGAAACCCCGCTGATTTGCAACATCATCTCTCAGGGTCCCTTCTGAG 589
QY 340 AGACTATGGAACACAGACGCTCTTATGTACAACAGCGCGGACGAAGGTTGACAGCAAG 399
Db 590 ATGATCTGATCCCGACATTTGTTCTTACAACAATGAGATGGGAGTTTGCAGTGACC 649
QY 400 TATCCACGAACGCTGGTGGTGGGACCAACGCGTCTGTACGTCGCGCGCGGATC 459
Db 650 CACATGACCAAGGCCACCTCTTCTCCACGGGCACTGTGCACCTGGGTGCCCGGCCATC 709
QY 460 TTCAAGAGCACTTCAAGATCGACATCACTCTGTTCCCTTCCAGACCAACAGATGCGAG 519
Db 710 TACAAGAGCTCTTGCAGCATCGAGCTCACTCTTCCCTTCCAGACGAGAACTGCAAG 769
QY 520 ATGAAGTTTGGCAGCTGGAATATGATGATGATGATGATGATGATGATGATGATGATG 579
Db 770 ATGAAGTTTGGCTCTCTGGAATATGATGATGATGATGATGATGATGATGATGATG 826
QY 580 GGGGGGGGAGATATAAGCAGTCTTGTCAACGAATGGGAATGGGAGTTAATAGGATCCCC 639
Db 827 CAGACTGTGACCTTGAGGACTACTGGAGAGCGGCGAGTGGGCCATCTGTCATATGCCAG 886
QY 640 GGCAAGCGCAACAGAGATCTACTACACTGTGTTCGGAGCCATACATTCGACATCACCTTT 699
Db 887 GGCACCTTACAACAGCAAGAGTACGACTGTCTGCGCGAGATCTACCCCGACGTCACCTAC 946
QY 700 GCGGTGGTGTATCCGGAGGAAAACGCTCTACTACTTCTTCTTCTTCTTCTTCTTCTTCT 759
Db 947 GCCTTGTCTATCCGGGGGCTGCGGCTCTTCTACACCATCAACCTCATCATCCCTGCTG 1006
QY 760 CTCACTCCCTCCATGGCTCTATTTGGGGTTCACCTTGCCTCCAGACTCCGAGAAAGTTG 819
Db 1007 CTCATCTCTGCTCTCACTGTGCTGGTCTTCTACTGCGCTCCGAGTGGCGGAGAGATC 1066
QY 820 TCTTTAGGTGTGACGATATTACTGTCTGACGCTGTCTTCAACATGGTGGGGAGACG 879
Db 1067 ACCTGTGCAATTTGGGTGCTGCTCACTACCGCTCTTCTGCTGCTCATCTGAGATC 1126
QY 880 ATGCCAGCAGCTCGGACGCGCTGCTGCTCGGACCTTACTTCAACATCATCATGATTC 939
Db 1127 ATCCCGTCCACCTCGCTGGTCTATCCCGCTCATCGGCGAGTACCTGCTGTTTCACTATC 1186
QY 940 ATGCTGGCTTCTCCGCTCGCTCTCCACCATCTCTCACTACCATCCAGCCGACGCA 999
```


[illegible]